PART I

INTERACTING ELECTRONS: BEYOND THE INDEPENDENT-PARTICLE PICTURE

1

The many-electron problem: introduction

> The calculation of a wavefunction took about two afternoons, and five wavefunctions were calculated in the whole . . . .
>
> Wigner and Seitz, 1933

Summary

In order to explain many important properties of materials and phenomena, it is necessary to go beyond independent-particle approximations and directly account for many-body effects that result from electronic interaction. The many-body problem is a major scientific challenge, but there has been great progress resulting from theoretical developments and advances in computation. This chapter is a short introduction to the interacting-electron problem, with some of the history that has led up to the concepts and methods described in this book.

The many-body interacting-electron problem ranks among the most fascinating and fruitful areas of research in physics and chemistry. It has a rich history, starting from the early days of quantum mechanics and continuing with new intellectual challenges and opportunities. The vitality of electronic structure theory arises in large part from the close connection with experiment and applications. It is spurred on by new discoveries and advances in techniques that probe the behavior of electrons in depth. In turn, theoretical concepts and calculations can now make predictions that suggest new experiments, as well as provide quantitative information that is difficult or not yet possible to measure experimentally.

This book is concerned with the effects of interactions between electrons beyond independent-particle approximations. Some phenomena cannot be explained by any independent-electron method, such as broadening and lifetime of excited states and two-particle bound states (excitons) that are crucial for optical properties of materials. There are many other examples, such as the van der Waals interaction between neutral molecules that arises from the dipole–induced dipole interaction. This force, which is entirely due to correlation between electrons, is an essential mechanism determining the functions of biological

systems. Other properties, such as thermodynamically stable magnetic phases, would not exist if there were no interactions between electrons; even though mean-field approximations can describe average effects, they do not account for fluctuations around the average. Ground-state properties, such as the equilibrium structures of molecules and solids, can be described by density functional theory (DFT) and the Kohn–Sham independent-particle equations. However, present approximations are often not sufficient, and for many properties the equations, when used in a straightforward way, do not give a proper description, even in principle. A satisfactory theory ultimately requires us to confront the problem of interacting, correlated electrons.

It is challenging to develop robust, quantitative theoretical methods for interacting, correlated electrons, but there has been great progress. The exponential growth of computational capability and the development of algorithms have turned theoretical ideas into practical methods. The field has gone a long way, from calculations for single atoms or the homogeneous electron gas, to applications that impact not only the traditional fields of solid-state physics, materials science and chemistry, but also diverse areas ranging from engineering to biology, from archeology to astrophysics. Progress in all these disciplines can be greatly enhanced by quantitative calculations based on the fundamental laws of quantum mechanics. In this chapter we will see how the bridge is made between theoretical foundations and computational realizations.

1.1 The electronic structure problem

What we want to describe is the behavior of atoms, molecules, and condensed matter, which is governed by the quantum statistical mechanics for electrons and nuclei interacting via the Coulomb potential. The essential ingredients[1] are contained in the hamiltonian

$$
\hat{H} = -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 - \sum_{i,I} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}
$$
$$
- \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|}, \tag{1.1}
$$

where electrons are denoted by lowercase subscripts and coordinates $\mathbf{r}_i$ and nuclei are denoted by uppercase subscripts, coordinates $\mathbf{R}_I$, charge $Z_I e$, and mass M_I. Throughout much of the book, we will use *atomic units* defined such that $\hbar = m_e = e = 4\pi/\epsilon_0 = 1$.

---

[1] The hamiltonian in Eq. (1.1) displays the *dominant* terms for the problems that are addressed throughout this book. Other terms are included when needed: externally applied electric and magnetic fields can be treated as an added scalar potential, a vector potential in the kinetic energy operator as $(i\nabla_j + \frac{e}{c}\mathbf{A})^2$, and Zeeman terms for the spin in a magnetic field. Time-dependent electromagnetic fields are included later in perturbation theory. Spin–orbit coupling and other relativistic effects are important in heavy atoms; however, they are one-particle terms that originate deep in the core and can be included in an ion core potential (or pseudopotential) as explained in [1, Chs. 10 and 11]. Corrections to the Coulomb potential due to quantum electrodynamics are small on the scale of interest for the chosen applications and are neglected.

Thus the unit of length is a Bohr ≈ 0.0529 nm and energy is a Hartree = 2 Rydberg $\approx$ 27.211 eV (or the equivalent in temperature of 315 775 K).

In Eq. (1.1) there is only one term that can be regarded as small: the nuclear kinetic energy, proportional to the inverse mass of the nuclei $1/M_I$. For the most part we ignore this term to concentrate on the electronic problem with fixed nuclei, using the Born–Oppenheimer (adiabatic) approximation [2]. Once we have neglected the kinetic energy of the nuclei and fixed their positions,[2] the final term, the interaction of nuclei with each other, is a constant that can be added to the zero of energy. Hence, the hamiltonian essential for the theory of interacting electrons consists of the first three terms, the kinetic energy of the electrons, the electron–nucleus interaction, and the electron–electron interaction,[3] which we can write as

$$\hat{H} = \hat{T}_e + \hat{V}_{en} + \hat{V}_{ee}. \tag{1.2}$$

This equation encapsulates the electronic structure problem: the first and last terms are universal for all problems, and the information specific to any system is contained in the middle term, a potential that acts equally on all electrons. There is a vast collection of approximations and techniques to deal with this hamiltonian because there are so many phenomena and materials, and because the two-body electron–electron interaction term $\hat{V}_{ee}$ makes the problem so difficult. However, in recent years it has become possible to compute many properties with methods derived directly from Eqs. (1.1) and (1.2). These methods and the resulting properties are the topic of this book.

1.2 Why is this problem hard?

Let us first address the question of why the straightforward solution of the many-body Schrödinger equation is so difficult. The underlying reason is the dimensionality of the problem, "the curse of dimensionality." A many-body wavefunction does not factorize because of the electron–electron interaction. In three dimensions and for N electrons it is a complex-valued function of $3N$ variables. Suppose we try to write the complete many-electron wavefunction for N electrons in a basis. With this aim, we introduce a single-particle basis set; there must be many basis functions to describe the wavefunction accurately. Let the total number of single-particle basis functions be M, with $M/N > 1$. Since electrons are indistinguishable fermions, the total basis for the many-body wavefunction consists of all Slater determinants of matrices with size $N \times N$ that can be constructed from the single-particle basis set. The number of these determinants is

---

[2] Quantum effects in the nuclear motion are important in some cases, e.g., hydrogen, helium, and other light nuclei. Those quantum effects can be treated directly, for example, by using path integral methods as described in Ch. 25 or approximately, for example, by assuming harmonic phonons. For heavier elements, the dynamics of the nuclei can be handled with the classical molecular dynamics or Monte Carlo methods.

[3] For bulk systems, there is a difficulty in dividing up the various terms that involve the Coulomb interaction because each term is individually infinite. This can be handled by adding a uniformly charged background to each such term, so each one is charge neutral.

$$\left(\begin{array}{c} M \\ N \end{array} \right) = \frac{M!}{N!(M-N)!} \approx e^{CN}, \tag{1.3}$$

where $C > 0$. This is typically a huge number; for example, consider two carbon atoms, with their $N = 12$ electrons. Suppose we represent the orbitals with only 36 basis functions. This small representation yields more than 10^9 determinants (see Ex. 6.1).

The wavefunction for each eigenstate is a linear combination of these determinants. Though there are often other symmetries, such as inversion or lattice translation, they do not help enough to reduce the size of this linear combination to a reasonable amount. It follows that to store a complete wavefunction and to manipulate it becomes exponentially difficult as the number of electrons becomes large. The straightforward application of methods such as the configuration interaction method [3], used in quantum chemistry to compute properties of small molecules by exact diagonalization of the hamiltonian matrix, is therefore limited to small numbers of electrons.

The difficulty in computing and storing one wavefunction (typically the ground state) is only the tip of the iceberg. For systems at a non-zero temperature, all states that are thermally occupied need to be computed and one often wants the response to external perturbations including spectra, thermodynamic functions, and more. To solve these problems requires computation not only of the ground state, but also of many excited states. The number of states required will scale exponentially with the system size. Finally, if we want to do a molecular dynamics simulation, the calculations must be made many millions of times with the forces on all of the atoms calculated at each step.

The path to deal with these issues is to realize that it is not essential to tabulate the complete many-body wavefunction. What is usually needed are single numbers, such as the energy or its derivatives, one-particle properties such as the electron density, and spectra derived from one- or two-particle, or higher-order, correlation functions. All these are expectation values. In Ch. 5 we introduce some of their most important properties. Much of the rest of this book is then devoted to alternative ways to approach the problem: methods to compute measurable properties without precise knowledge of the many-body wavefunctions.

One way this can be done is in terms of Green's functions, introduced in Ch. 5, which represent directly the experimentally measurable spectra. These are developed in Part III, along with examples of applications. Another approach is through techniques that directly simulate the many-electron system. Instead of trying to represent explicitly the complete many-body wavefunction, it is sampled instead. These methods, called quantum Monte Carlo (QMC) methods, are discussed in Part IV.

There is no universal approach to find feasible, well-founded methods to deal with the many-body problem. There are many different materials, many phenomena, and many different criteria to obtain the accuracy needed for a useful result. For some problems, such as the relative stability of dissimilar structures that may be separated by very small differences in energy, we might require the accuracy in the energy difference to be much less than the ambient temperature, 300 K $\approx$ 0.026 eV $\approx 9.5 \times 10^{-4}$ a.u., which is often called "chemical accuracy." In other cases, one may find it satisfactory to calculate a bandgap to

say 0.1 eV, or to detect the presence or absence of a satellite in an energy window of 1 eV. The nature of the problem and the required precision, together with the available computer power, will finally determine the appropriate methods to use.

1.3 Why is the independent-electron picture so successful?

Underlying all independent-particle approaches is the idea that each electron interacts with an effective potential that mimics to some extent the effects of the other electrons. A single-particle wavefunction can be expressed in a basis of dimension M, the same size as assumed in the many-body problem in the previous section. However, there is a great simplification since an independent-electron state needs only one determinant. One can work with wavefunction methods that require only $M \times M$ matrix operations that scale as M^3 or better, as opposed to the exponential scaling of the interacting many-body problem.

The independent-particle picture is extremely successful: it is used to classify solids into metals, semiconductors, and insulators, and to establish much of the present-day understanding of solids.[4] Quantitative calculations were put into practice more than 80 years ago, for example, numerical mean-field calculations for atoms in 1928 by the Hartrees [5].[5] Today, essentially all practical, quantitative many-body methods depend on information from independent-particle calculations as input, and the analysis is often in terms of single-body concepts. It is important to recognize the reasons for the success of independent-particle concepts and techniques. These provide invaluable lessons for the much more difficult problems encountered in many-body theory, and they reveal the cases where it is essential to go beyond independent-particle pictures.

Independent fermions

Any reasonable independent-particle approach recognizes the fact that electrons "see" each other, since an orbital can only be occupied by one electron of a given spin; by "independent" we mean "independent except for the requirement of the exclusion principle." Indeed, the very origin of the Fermi–Dirac statistics and the term "fermion" is an example of the power of independent-particle arguments. In his 1926 paper [7],[6] Fermi arrived at the conclusion "it is required to admit that an electronic orbit is already 'occupied' when it contains only one electron" by considering electrons in heavy atoms, where he argued that the nuclear potential dominates over electron–electron interactions. From the observation that

---

[4] Further background and references can be found in [1] and an exposition of the history of the theory of solids can be found in the book [4].

[5] D. R. Hartree was aided by his father W. R. Hartree, a businessman with an interest in mathematics, who carried out calculations on a desk calculator. Together they published numerous calculations on atoms. D. R. went on to become one of the pioneers of computer science and the use of electronic computers, and he published a book on the calculation of the electronic structure of atoms [6].

[6] The paper by Fermi (see [8] for a translation) refers to earlier work by Stoner [9] on "The distribution of electrons among atomic levels" and Pauli [10] on the interpretation of atomic spectra. Apparently the development of the determinant formulation by Dirac [11] in 1926 was independent, though he refers to Fermi.

the lowest *s* state holds only two electrons, Fermi drew sweeping conclusions on the statistics of particles in an ideal monatomic gas. This was recognized by Dirac in his 1930 book on quantum mechanics [12], where he says ". . . . a special statistics, which was first studied by Fermi, so we shall call particles for which only antisymmetric states occur in nature *fermions.*" In the independent-particle approximations, the antisymmetry of the full many-body wavefunction is hidden, but it is implicitly present through the use of Fermi–Dirac statistics.

Symmetries and conservation laws

The antisymmetry of the wavefunctions is a general feature of systems of fermions, as is time reversal in the absence of external magnetic fields. In specific systems the hamiltonian may have additional symmetries, such as inversion and rotational invariance, and continuous or discrete translational symmetry in a homogeneous gas or a crystal, respectively. Coming together with a symmetry there is a conservation law, e.g., conservation of momentum for translation symmetry. It may be detrimental for a theory to violate these conservation laws, and, on the contrary, a key for success is to take them into account. In independent-particle methods each state of a many-particle system can be specified by the occupation numbers of the independent-particle states, and it is not difficult to take the symmetries into account. In a many-body theory that treats the interaction between particles, these states are mixed and the problem is not so simple.

Two examples show the issues that one may face in many-body calculations. In cases where there is a discrete quantum number, such as plus or minus parity, the many-body states can be readily constructed to have the correct symmetry. So long as the energy levels are discrete, the wavefunction and energy for each state can be found by exact diagonalization, at least in principle. Nevertheless, it is very useful to identify independent-particle states that are as close as possible to the actual states to facilitate the calculation and to interpret the results. However, in condensed matter the conserved quantity can take a continuous range of values, namely momentum that can take any value in a homogeneous system or "crystal momentum" that varies continuously in the Brillouin zone. For a given energy there can be an infinite number of ways the momentum can be shared among the electrons. In an independent-particle picture each of the possible states with energy E and momentum $\mathbf{K}$ consists of independent excitations with momenta $\mathbf{k}_1, \mathbf{k}_2, \ldots$ where $\mathbf{k}_1 + \mathbf{k}_2 + \cdots = \mathbf{K}$ and $\varepsilon_1 + \varepsilon_2 + \cdots = E$. In a many-body theory, however, these states are mixed by the interaction and it is not always easy to develop approximations that are guaranteed to obey the symmetry. In fact, a fundamental paper for many-body Green's function methods by Baym [13] is devoted to the conditions for "conserving approximations" that are described in Sec. 8.5. Again, independent-particle pictures can provide guidance. For example, the Bloch theorem applied to independent particles leads to the band structure for a crystal $\varepsilon_\mathbf{k}$; as brought out in Sec. 2.4 and other places in this book, the spectral function including interactions can often be understood as a broadened version of the band structure.

Fermi liquid theory and continuity

Perhaps the most direct use of independent-particle concepts to characterize interacting systems is Fermi liquid theory, proposed by Landau [14–16] and summarized in Sec. 3.4. This theory, originally proposed for liquid ^{3}He, assumes the low-temperature properties are determined entirely by free-particle-like low-energy excitations near the Fermi surface, called quasi-particles, with renormalized masses and weak effective interactions. It is not obvious that excitations in liquid ^{3}He should behave like a weakly interacting gas; the same reasoning applies to electrons in solids, where the magnitude of the interaction ($\approx e^2/R \approx$ several eV, with R a typical atomic size) is not small.

According to Alexey Abrikosov,[7] Landau arrived at his proposal by asking himself, "What is conserved?" The fact that there are fundamental features such as conservation laws unchanged by the interaction suggests continuity. Suppose that we could turn on the electron–electron interaction continuously.[8] Then there should be a one-to-one correspondence between the excitations of the non-interacting system and the real excitations. Although the momentum is shared among electrons, the total current is the same as for non-interacting particles since the total momentum is conserved. The reasoning should apply so long as the system evolves continuously, i.e., there is no phase transition as the interaction is turned on. Continuity is a critical part of the reasoning regarding the Fermi surface and bandgaps in interacting systems (e.g., Chs. 13 and 16–21).

The Fermi surface is determined by a step in the momentum distribution at zero temperature. The existence of such a step, as well as the shape of the surface in **k**-space where the step occurs, are straightforward to understand for independent electrons, but less obvious in the interacting case, because of the electron–electron scattering. Both features are *assumed* in Fermi liquid theory, but in order to determine them, one must resort to other methods, which is not so easy in a many-body theory. Fortunately, the "Luttinger theorem" (Sec. 3.6) states that, so long as the actual interacting system can be considered to evolve continuously from a non-interacting system, the volume enclosed by the Fermi surface does not change. This is sufficient to determine the surface in liquid ^{3}He and the electron gas, since the surface is spherical and is determined by the volume. The problem is not as simple in a crystal, but the conservation of the volume is a welcome guidepost for many-body calculations.

The situation changes if the connection is broken. For bulk systems, discontinuities in properties can only occur at phase transitions. Phases can be classified according to symmetry and described by order parameters, as formulated by Landau [18, 19]. More recent developments on topological order are providing qualitatively new classifications and phenomena (see, e.g., Sec. 6.3 and [20, 21]). Although the continuous connection to one and the same independent-particle theory may be lost, nevertheless the continuity arguments can still be applied so long as one is careful to connect to an independent-particle theory with an effective potential with the same symmetry as the actual phase. To determine the

---

[7] Private discussions in 2005.
[8] This can now be done in cold atom experiments using the Feshbach resonance [17].

stable phase and the transition points is therefore one of the main challenges for many-body calculations.

1.4 Development of theoretical approaches to the many-body problem

When is the independent-particle picture not sufficient?

There are many examples of cases where the independent-particle picture clearly meets its limits. Some phase transitions and ordered states would never occur if the electrons did not interact. One is the Wigner crystal transition [22] (Sec. 3.1), where the electrons break translation symmetry at low density. Other examples are ordered magnetic states: there is no reason for spins to order antiferromagnetically without an electron–electron interaction. In Ch. 2 we will see several examples of failures of the independent-particle picture. In each case an independent-particle or to be precise, *one single* independent-particle description does not succeed in capturing the essential effects.

A different type of example is the density of states at the Fermi energy. In Bardeen's thesis as a graduate student with Wigner, he showed that the Hartree–Fock approximation introduces a singularity at the Fermi surface [23]. This conclusion applies to all metals since it depends only on analysis of an integral where the integrand has a singularity due to the long-range Coulomb interaction (see also Sec. 11.3). The observed smooth dispersion of excitations can be explained only if the interaction is screened, i.e., correlation of electrons that reduces the effective interactions. Screening is an essential feature of the methods in this book. The Green's function methods are developed in terms of a screened Coulomb interaction called W in the GW method (Chs. 11–13) and the screened effective interaction U in applications of dynamical mean-field theory (Chs. 19–21). Screening is also a key ingredient in the variational Monte Carlo method (VMC, Ch. 23), where the correlation (Jastrow) factor is often derived from the random phase approximation (RPA): to lowest order in the electron–electron interaction it corrects the Hartree–Fock wavefunction for effects of the interaction. All these developments go beyond a static mean-field description.

Genesis of theoretical methods for many-body calculations

Many methods in use today can be traced back to their roots more than 50 years ago. By 1930 there were accurate calculations by Hylleraas [24] of the ground-state energies of two-electron atoms and ions. In condensed matter physics, quantitative calculations on sodium were reported by Wigner and Seitz in 1933 and 1934 [25, 26] (see also [22]). The first paper [25] provided the picture of correlation in which one electron in a Wigner–Seitz cell prevents another electron from occupying the same cell. In their 1934 paper [26] they explained the effects in terms of the "Fock picture," called "one-electron," and the correction that lowers the energy. As phrased by Wigner [22], "This energy will be called the 'correlation energy'." Of course, this is hard to calculate and they proposed an interpolation between the low- and high-density limits that compares very well with the
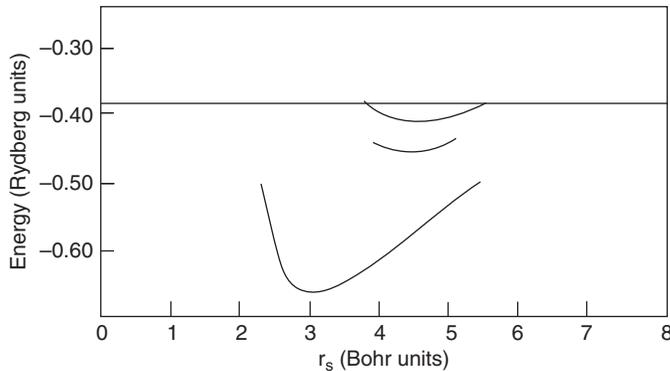
Figure 1.1. Total energy of sodium as a function of the Wigner–Seitz radius: reproduction of Fig. 4 from the 1934 paper by Wigner and Seitz [26]. The bottom curve is the energy of the lowest state in the band, and the top curve is the total energy in the Hartree–Fock approximation. The middle curve (from Fig. 6 of that paper) is the total energy including the estimated correlation energy. The horizontal line is the calculated atomic energy [26]. We can see that correlation increases the binding energy by more than a factor of 2; in fact, the result is within 15% of the measured binding energy.

best-known calculations of the correlation energy using diffusion Monte Carlo (Ch. 24) (see, e.g., [1, Ch. 5]). The results for Na are reproduced in Fig. 1.1, showing that the binding energy is greatly increased by correlation. Remarkably, the final result for the binding energy is close to the measured value.

Many of the most important developments in theoretical methods occurred in the late 1940s and 1950s. Two notable developments were the invention of path-integral methods [27] and Feynman diagrams [28]. With the advent of modern computers, the former has become an important non-perturbative method for studying correlated particles. The latter provides the graphical interpretation that has brought to life the summation of diagrams that are the basis for an untold number of quantum many-body techniques. In the same period was the advent of the Dyson equation and the functional formulation of the quantum many-body problem by Schwinger, Tomanaga, and others.[9]

Essential features of screening in a metal were captured by the RPA of Bohm and Pines in 1951–53 [30], which leads to the Lindhard dielectric function [31] for the electron gas. In 1958, Hubbard [32] showed that the RPA follows from the diagrammatic bubble expansion for the dielectric function and in 1959, Ehrenreich and Cohen [33] showed that it was tightly linked to the time-dependent Hartree approximation.

The late 1950s and early 1960s witnessed developments in diagrammatic techniques, including the work of Gell-Mann and Brueckner in 1957 [34], Galitskii and Migdal in 1958 [35], and Luttinger and Ward [36, 37], among many others. The book by Abrikosov, Gorkov, and Dzyaloshinski [38] in 1963 is remarkable for elucidating the physics and

---

[9] The 1949 paper by Dyson [29] gives references to Feynman, Schwinger, and Tomanaga.

in-depth coverage of the latest major development at the time. Baym and Kadanoff [13, 39] and earlier work of Martin and Schwinger [40] provided a general framework for characterizing functionals of the Green's functions and a concise statement of conditions for conserving approximations.

The years 1964–65 were especially significant in the development of the most widely used methods today. Density functional theory (DFT) by Kohn, Hohenberg, and Sham [41, 42] came after the more general functionals of Luttinger and others, but it introduced new ideas that led to practical methods. It was a major turning point that, with the advent of powerful computers, revolutionized computational materials modeling. A further landmark is Hedin's 1965 paper [43]. The abstract says "there is not much new in principle in this paper," but that work identified aspects of the many-body problem that are especially relevant for excitations in solids, and it established the "GW" approximation. McMillan (while a student of Bardeen) carried out variational QMC simulations of ^{4}He [44], the prototype for modern calculations, by realizing the relation between quantum many-body theory and the classical many-body computer simulation techniques that had been developed in the previous decade. The period 1963–65 witnessed the papers by Gutzwiller and Hubbard (see Sec. 3.2) that defined the models for interacting electrons on lattices and provided much of the conceptual structure for methods such as dynamical mean-field theory (DMFT). Together, the acronyms DFT, GW, QMC, and DMFT encompass most of the first-principles computations for materials today.

1.5 The many-body problem and computation

Moore's law (that the number of transistors per chip doubles roughly every year) has changed science in a profound way. Theory is no longer only a "pencil and paper" activity and we are no longer limited to the calculation of a wavefunction that requires two afternoons, as described in the quote from Wigner and Seitz at the beginning of this chapter. For electronic structure theory the advent of computers has been a game-changer, with the capability for quantitative calculations of properties of materials and the development of new approaches not envisioned in the early days of quantum mechanics before the invention of computing machines.

The combination of theory and numerical calculations allows us to get closer and closer to reality, with fewer approximations and more understanding. Quantitative calculations within independent-particle approximations using computers started in the 1950s, but it was only with the computers of the 1970s that calculations could be done with a sufficiently large basis that different methods could give the same answer – thus moving from debates about numerical approximations to addressing the deeper physics: how to deal with the interacting-electron problem. Several aspects highlight this important evolution:

- The great advances in the theory of interacting-electron systems in the 1950s and 1960s were applied to model systems such as the homogeneous electron gas. But it was only in the 1970s that computers and algorithms made possible QMC (Chs. 22–25) calculations