INTRODUCTION TO MODERN LIQUID STATE THEORY: 
FROM SIMPLE MODELS TO APPLICATIONS 

PART I: BASICS 

Ivo Nezbeda 

Ústí nad Labem 
2014
Title: Introduction to modern liquid state theory: From simple models to applications
Part I: Basics

Author: prof. RNDr. Ivo Nezbeda, DrSc.

© Publisher: J. E. Purkyně University in Ústí n. Labem, Faculty of the Environment

This publication was supported by the OP Education for Competitiveness project: EnviMod – Modernization of education in technical and natural sciences at UJEP with respect to environmental protection.
Reg. No.: CZ.1.07/2.2.00/28.0205

Free copy
## Contents

**Preface**  

1 **Introduction**  

2 **Systems of interacting particles: Basic considerations**  
   2.1 Simplifying assumptions and basic relations  
   2.2 Pairwise additivity  
   2.3 Classes of fluids  
   2.4 Simple fluid interaction models  
   2.5 Realistic potential models  

3 **Systems of interacting particles: Correlation functions**  
   3.1 Molecular correlation functions  
   3.2 Partial correlation functions  

4 **Pseudoexperimental techniques**  
   4.1 Basics of molecular dynamics  
      4.1.1 General remarks  
      4.1.2 Verlet method  
      4.1.3 Predictor-corrector algorithms  
      4.1.4 Remarks and comments  
   4.2 Basics of Monte Carlo method  
      4.2.1 Naive versus smart sampling  
      4.2.2 Metropolis algorithm  

5 **Basic theoretical tools**  
   5.1 Functional differentiation  
   5.2 Diagrammatic technique  
   5.3 Integral equations  
   5.4 Perturbation theories  

A **Appendix: Basic thermodynamic relations**
PREFACE

The study of fluids (i.e., liquids and dense gases) at the molecular level, a multidisciplinary endeavour pursued by physicists, mathematicians, chemists, chemical engineers, and (most recently) biologists, is an area of relatively modern study and progress. The first significant results date back to the end of the 19th century and are associated primarily with the work of van der Waals and the Dutch School. It was only then that the view that matter is made up of molecules started gaining ground and that viewpoint was fully accepted only after quantum theory was developed in the first decades of the last century. Although this laid the foundations for subsequent study, no significant progress took place until the middle of the last century.

Although basic molecular theories of solids and gases were developed in the first half of the 20th century, similar theories governing the behavior of liquids remained unknown. It wasn’t that liquids were ignored or out of the focus of theoreticians; a number of methods to deal with liquids were developed, including cluster expansions [1], integral and integro-differential equations for correlation functions [2] and lattice/cell theories [4]. A key historical marker was the 1951 book by John Barker, which sounded the death-knell for pursuing the study of liquids as being primarily solid-like. Use of the first two approaches were hindered by the lack of computational machinery to implement them. Progress was also hindered for the simple reason that liquids represent a system of many interacting particles and their description thus requires, in addition to a statistical mechanical theory, also a Hamiltonian describing the underlying molecular interactions. Since neither of these ingredients were available (and have not been until very recently) in an accurate form, it was not possible to attribute the lack of progress to deficiencies in either the interaction model or the theory.

The aforesaid situation was also reflected by textbooks and monographs on statistical mechanics of the time [5, 6]. Although they contained chapters on the ideal gas and the ideal crystalline solid, the Fermi and Bose gases, they contained possibly only a chapter dealing with real gases at the level of the van der Waals equation and the virial expansion. They contained not a word concerning dense liquids. The state-of-the-art molecular view of liquids
themselves at the time was summarized in a comprehensive monograph by Hirschfelder et al. [7].

A real breakthrough took place in the 1960's, made possible due to advances in computational technologies (digital computers) and the development of new computational methods particularly suited to them: Monte Carlo [8] and molecular dynamics techniques [9, 10]. These required only the specification of a Hamiltonian to produce essentially numerically exact (within a certain precision) results for a fluid described by it. On the one hand, molecular simulations thus provided ‘exact’ data to test the accuracy of theoretical predictions and could be viewed as computer experiments on virtual real substances; on the other hand, comparison of the simulation data with experimental results could assess the appropriateness of the assumed Hamiltonian. Thus, in the following decades, the field of research of liquids at the molecular level flourished, and a rapid development took place.

The computer simulations both generated quantitative results for problems defying analytical solutions and provided theoreticians with a wealth of data with which to test theories. They also provided an impetus for quantum chemists to develop improved Hamiltonians for particular systems. These scientific motivations were augmented by practical demands from increasingly sophisticated industry, in view of the fact that the overwhelming majority of chemical reactions take place in the fluid phase and, consequently, almost 80 per cent of the activities of chemical industry deal with separation processes to extract final manufactured products from fluid mixtures.

The new developments were also reflected by textbooks on statistical physics published in the last decades of the 20th century and at the turn of the century. In addition to the description and development of the basic tools used to treat systems of many interacting particles they also included chapter(s) dealing with liquids e.g., [11, 12, 13]. Moreover, a relatively large number of monographs devoted specifically to molecular theories of liquids also emerged (e.g., [14, 15, 16, 17, 18, 20]). Regardless of the target reader audience, they contain similar content. After a short discussion on intermolecular interaction models, they focus on integro-differential equations and, primarily, integral equations (correlation function theories), along with their theoretical basis and properties, and the derivation of various approximations to make them (analytically) solvable; finally, perturbation theories are also discussed. All these methods are explained primarily for simple fluids, that is, the interaction models assume pairwise additivity of the configurational energy in terms of an interaction potential $u$ that depends only on the separation distance $r$ between molecules, $u = u(r)$. Practically only one monograph covered systems with orientation-dependent interaction models, which dealt with non-spherical hard bodies and simple models with point multipoles [21]. Monographs on theories of liquids continue to emerge today, with more detailed descriptions of the theories and methods for their implementation. Again, they consider only simpler models, and they differ from each other mainly in
the selection of applications, see, e.g., [22, 23] The most recent frontier of computer simulation approaches involves so-called *ab-initio* simulation methods that calculate the Hamiltonian approximately by numerically solving the Schrödinger equation (typically using a form of density functional theory) at each time step of a molecular dynamics simulation, thereby obviating the need for an approximate force field. Although these approaches are currently very computationally expensive, they will likely see increasing future improvements and use.

From the beginning of this century another very important *qualitative* change in the course of development in the molecular-level approach to fluids (which is now increasingly referred to as “soft matter”) has been observed. This is again associated with the development of molecular simulation methodology. On the one hand, the power of computers has been increasing at an unprecedented pace, making it possible to study quite complex systems. In addition, sophisticated simulation packages have become available, either commercially or as freeware. This means that possibilities to study liquids at the molecular level has been opened up to a large community of non-specialists; in other words, the main focus has shifted now from theoretical statistical mechanical considerations to the use of simulations as a primary investigative tool. The scientific problems have thereby practically shrunk to the choice of the interaction model; the rest is taken care of by the software.

To understand why the spectrum of researchers using the molecular-level approach to describe fluid behaviour has grown significantly, it is important to realize that a primary output of molecular simulations (and target of theories) have typically been thermophysical properties (equation of state, heat capacity, viscosity, etc.). These are used in turn to both refine macroscopic theories of the systems, and to refine their the descriptions of the underlying intermolecular interactions. They also provide feedback for quantum chemists to improve the quality of the Hamiltonians used in the simulations. However, in some areas of science only a *qualitative* molecular picture is required to replace ad hoc speculations, and this is facilitated by the availability of powerful computers and software packages. The molecular simulation approach has become thus commonplace in many other fields, including mechanical engineering and material design, biology, and even medicine, often shedding light on problems that are inaccessible by other methods. As examples we may mention the nonequilibrium phenomena of the rheology of lubricants [24, 25], protein self-assembly [26] (potentially related to Alzheimer’s and Parkinson’s diseases), and molecular insight into industrial processes as, *e.g.*, the needle-less electrospinning process used to produce nanofibers from polymer solutions [27].

The opening up of the field of molecular approach to a larger user audience is also accompanied by a danger. To carry out molecular simulations using available software packages does not
automatically mean that the obtained results must correspond to reality. Along the path from setting an intermolecular potential model, to choosing the appropriate simulation methodology, specifying a number of technical simulation parameters, and analyzing the results, there are a number of pitfalls (see, \textit{e.g.}, an interesting collection of warnings by Frenkel [28]).

Since fluid systems are practically ubiquitous, the community dealing with fluid systems is enormous. In addition to specialists/theoreticians there are two large and important user groups with quite different goals who are making use of molecular-level computer simulation technology. One is represented by (primarily chemical) engineers who, for efficient equipment design need descriptions of fluid behavior which can be either expressed directly by means of macroscopic theoretically based equations in analytical form or by means of equations capable of rapid numerical solution. An important component of these expressions is a set of parameters, which must be determined to adequately predict known experimental results. Prior to their use in equipment design, the equations are typically tested by comparison with both the computer simulation results and with experimental data for the systems of interest. Consequently, in addition to their intrinsic scientific interest, molecular-level computer simulations can find use in providing pseudo-experimental data with which to fit parameters in macroscopic expressions. This approach can be much more cost-effective than performing expensive real experiments to provide the often extensive database with which to fit the multiple parameters in the empirical models. It is also interesting to note that, similarly to the aforementioned development of molecular-level computer simulation packages in the scientific community, a parallel development has occurred in this user community in recent decades, with the availability of commercial chemical process simulation software such as AspenPlus, HySim, Prosim, ILL, etc that incorporates an array of macroscopic expressions to describe fluid properties of interest. The vendors of these packages regularly scan the literature to incorporate the latest theoretical developments.

The other large user group is represented by biochemists, who typically need to describe big flexible molecules in aqueous environments. Unlike the case of the engineers, due to the complexity of the systems involved and the consequent unavailability of theoretically-based equations in analytical form, this user group depends more directly on the use of molecular-level simulation methodology to describe systems within their domain of interest. To this end, they need to know how to formulate the molecular interaction model, simplifying it sufficiently to enable some conclusions to be drawn on the relation between the molecular picture and the macroscopic observations.

The present textbook is intended to serve primarily the above mentioned communities of applied researchers whose goal is not to develop a new theory of fluids but need the results for
their further research. The book is divided into two parts, the first presenting the basic principles and simulation algorithms, and the second part deals with applications. After a concise summary of the relevant statistical mechanical relations, Chapter 2 focuses systematically on the intermolecular interactions and their models, starting from simple fluids to modern site-site models describing more complex fluids. Chapter 3 is devoted to molecular correlation functions and their properties and use for obtaining experimentally observable macroscopic properties. Because there exist a large number of texts on molecular simulations, in Chapter 4 only the basic principles of molecular simulation methodology are described. The basic theoretical tools used to treat liquids are introduced in Chapter 5: diagrammatic techniques, cluster expansion, integral equations, and perturbation expansions. Complex mathematical derivations are omitted, emphasizing only the main ideas and their logical development. Whereas the content of Chapters 4 and 5 can be found in variety of available textbooks on theories of liquids, Part II, i.e., Chapters 6 through 11 deal systematically with specific classes of liquids, starting from the simplest model fluids up to aqueous solutions. Available results are assessed and explicitly given with potential recommendations and/or warnings.

This book has grown out from the advanced courses on Applied Statistical Mechanics and Theory of Liquids which I taught at Charles University and the Prague Institute of Chemical Technology and later at J. E. Purkinje University in Usti nad Labem along with the course Molecular Theory of Matter. It is assumed that the reader is familiar with the basic laws of thermodynamics, quantum mechanics, and statistical mechanics. It is also assumed that a molecular-level computer simulation package is available.

Horní Svetlá 2014
Chapter 1

Introduction

Everyday experience teaches us that there are three states of matter: gas, liquid, and solid. We can measure their temperature or volume, their response to an applied force (deformation) etc. This is the macroscopic point of view and the appropriate science to deal with these phenomena is thermodynamics and mechanics of continua. From this point of view a liquid is characterized by two basic features: it is fluid (in contrast to a solid it flows) and it is condensed (in contrast to a gas, which expands to the entire available volume, liquids fill only a part of the container). Liquid seems thus to form an interchain between gas and solid and this is also reflected by the fact that liquid can exist only within a certain range of thermodynamic conditions. This feature is well illustrated by the $P - T$ (pressure–temperature) of a common compound shown in Fig. 1.1.

Below the triple point pressure the matter exists either in a dense solid phase or in a very dilute gas phase. In addition to these stable phases, liquid may also be supercooled into a metastable phase until it reaches the state where it looses the mechanical stability (it would expand upon compression); the line formed by these states is called spinodal. The existence of the critical point makes it possible to change continuously, i.e., without a phase transition, gas into a liquid phase, see the $P - T$ phase diagram. These two states of matter, liquid and gas, are therefore referred to collectively as fluid. Finally, a fluid above the critical pressure but below the critical temperature is called dense gas.

From the microscopic point of view the matter is made up of particles (ions, atoms, molecules, etc.) and the appropriate science to deal with many-body systems is statistical mechanics. When we open a standard textbook on statistical mechanics published in the second half of the last century we found, after defining the basic notions and developing basic tools, the chapter on an ideal gas and very likely also on real gas, then the chapter on an ideal crystal and, perhaps, on electron gas as an example of a quantum system, and their applications. Unfortunately,

\footnote{It has become common nowadays to distinguish only two phases, ordered (crystal) and disordered (soft matter) phases.}
with only a few exceptions (e.g., an excellent and comprehensive textbook by McQuarrie [11] or a textbook by Chandler [13]) it also very often ends here. When liquids are mentioned then usually only in the context of the van der Waals equation but no chapter on theories of liquids and the evident question is: Why?

One reason can be immediately deduced from the previous paragraph: there is no word on an ideal liquid, i.e., a prototype model which would exhibit the main features common to all (or at least most) liquids. To understand why, we summarize in Table 1.1 the basic features of the three phases considered. Gas is characterized by low values of density which in the limit to zero leads to an ideal gas. Density may thus play the role of a small parameter and an expansion in its powers makes it then possible to estimate the properties of gas at finite densities by evaluating corrections to an ideal gas. With increasing density, however, more and more terms would be necessary until the expansion ceases to converge at all. This concept cannot be thus extended to liquids. At the other side of density values, the particles are so densely packed that particles around a randomly chosen particle form a cage making it impossible for this particle to sneak out. Particles are thus trapped in such cages only within which they can move. An idealization of this situation is the notion of an ideal crystal, i.e., a regularly arranged net of points around which the particles vibrate. Attempts to use the concept of a certain regular arrangement of particles with void lattice points to liquids lead to lattice (cell) theories [4] but also without
any considerable success (with possibly one exception, polymer systems). As we see, liquids do not make it possible to introduce *intuitively* any simplification/idealization and hence the term ‘ideal liquid’, i.e., a simple system which would capture the main features of the liquid phase, does not exist.

<table>
<thead>
<tr>
<th>Tabulka 1.1: Three states of matter: Basic facts</th>
</tr>
</thead>
<tbody>
<tr>
<td>gas</td>
</tr>
<tr>
<td>low density</td>
</tr>
<tr>
<td>⇓</td>
</tr>
<tr>
<td>molecules are, in average, far apart</td>
</tr>
<tr>
<td>⇓</td>
</tr>
<tr>
<td>their mutual interaction is very small (negligible)</td>
</tr>
<tr>
<td>⇓ (idealization)</td>
</tr>
<tr>
<td>interaction energy is zero</td>
</tr>
<tr>
<td>⇓</td>
</tr>
<tr>
<td>model: <strong>IDEAL GAS</strong></td>
</tr>
</tbody>
</table>

The other reason, surprisingly not often mentioned and acknowledged, is the fact that results of statistical mechanics, unlike other branches of theoretical physics, cannot be directly verified experimentally. The process of statistical mechanical calculations is compared in Table 1.2 with the laboratory experiment. The problem is that in these calculations there are involved two uncertainties: (i) an input model (Hamiltonian) estimating the intermolecular interactions in the system and supplied usually by quantum chemists, and (ii) an approximate statistical mechanical method. How a (better) theory can be developed when it is not clear whether the observed discrepancies between theory and experiment are due to crude theoretical approximations or due to an inaccurate Hamiltonian? Or both?

This situation lasted till the early of 1960’s when molecular simulations (computer experiments or, better, pseudoexperiments) [29, 30], along with availability of reasonably powerful digital computers, made it possible to get accurate (within experimental uncertainty) results for the given input model eliminating thus one uncertainty in the input definition of the problem, see Table 1.3. In other words, it has become thus possible to check accuracy of both theory (theory vs. simulation) and the input model Hamiltonian (simulation vs. real experiment).

Moreover, extensive simulations, first on simple models and later on fluids made up of gradually more and more complex molecules, revealed and for the first time also confirmed, relations
Tabulka 1.2: Comparison of the process of laboratory experiments and theoretical calculations

<table>
<thead>
<tr>
<th>Laboratory experiment</th>
<th>Theoretical calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>real compound in test tube</td>
<td>model compound in computer</td>
</tr>
<tr>
<td>⇄ (APPROXIMATE Hamiltonian)</td>
<td></td>
</tr>
<tr>
<td>⇓</td>
<td></td>
</tr>
<tr>
<td>carry out</td>
<td>apply APPROXIMATE</td>
</tr>
<tr>
<td>exptl. measurements</td>
<td>theory</td>
</tr>
<tr>
<td>⇓</td>
<td></td>
</tr>
<tr>
<td>experimental result</td>
<td>← ? compare ? → theoretical result</td>
</tr>
</tbody>
</table>

Tabulka 1.3: Comparison of the process of computer experiments and theoretical calculations

<table>
<thead>
<tr>
<th>Molecular simulation</th>
<th>Theoretical calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>model compound (Hamiltonian)</td>
<td>model compound (Hamiltonian)</td>
</tr>
<tr>
<td>⇓</td>
<td></td>
</tr>
<tr>
<td>carry out computer</td>
<td>apply APPROXIMATE</td>
</tr>
<tr>
<td>pseudo experiment</td>
<td>theory</td>
</tr>
<tr>
<td>⇓</td>
<td></td>
</tr>
<tr>
<td>experimental result</td>
<td>← ? compare ? → theoretical result</td>
</tr>
</tbody>
</table>

between the forces acting between the molecules and the observed properties of liquids and hence the basic mechanisms governing their behavior. As a feedback, this development enabled, on one side, to start developing more faithful and realistic and hence also more accurate intermolecular potential models and, on the other side, also to fully justify the use of very simple models. This was evidently also an impetus for theory that could be, step by step, improved. The successful theoretical developments include integral equations for molecular distribution functions and perturbation theories leading ultimately to analytic expressions for the thermodynamic properties of not only simple models such as the fluids of hard bodies but also for rather complex fluids and their mixtures. All this development has been motivated not only by scientific demands to improve the existing knowledge, but also by practical demands from increasingly sophisticated industry. Hence, developments in molecular theory are serving as a foundation for engineering design calculations in a growing number of industrial applications.
Chapter 2

Systems of interacting particles:
Basic considerations

2.1 Simplifying assumptions and basic relations

The ultimate goal of the statistical mechanics of matter is to provide methods of explaining and predicting the experimentally measurable quantities of a given substance in terms of the properties of its elementary constituent particles. The question which immediately arises is: What are the elementary particles that one should start with?

1. The fundamental particles are electrons and nuclei (e.g., the hydrogen’s and oxygen’s nucleus in the case of water).

2. The fundamental particles are individual atoms or the corresponding ions.

3. The fundamental particles are single molecules.

4. Accepting the fact that molecules can form (long-living) clusters, one can view fluid as a multicomponent system (e.g., clusters of hydrogen-bonded molecules in the case of water).

Each of the above listed possibilities can be used as a basis for an exact theory of fluids. In each case, however, the mathematical description of the system will be different and the inevitable approximations that we are compelled to use in any approach will have different physical significance. In making our choice among the various approaches we must maintain the balance between the more fundamental approach (usually more difficult to pursue) and the cruder, yet more feasible approach that leads to some useful information on the nature of the system.

The approach adopted in this book, and which corresponds to the modern statistical mechanics of fluids as it began developing in the early 1960’s, views fluids as a collection of individual molecules (level 3 above). Furthermore, when arguing at a general level, we will have in mind
small or only moderately big molecules as, e.g., those of lower hydrocarbons. The reason is not only simplicity. First, only after a successful theory for simple systems has been developed and molecular mechanisms governing their behavior have been understood it makes sense to move to complex systems. This level of understanding may serve as a guide and basis for developing a theory. Second, big molecules are made up of atoms or segments (collection of atoms) and knowledge of their (i.e., of atoms) behavior may be successfully made use of. Systems made up of large flexible molecules (higher hydrocarbons, polymers, biomolecules, etc.) will be only briefly addressed later in the second part of this textbook.

The fundamental relation connecting statistical mechanics (microworld), i.e., a system of \( N \) particles at the given \( T \) and \( V \) (canonical ensemble), with the thermodynamics (macroworld) is that between the Helmholtz free energy, \( A \), and the canonical partition function, \( Z \):

\[
\beta A = - \ln Z(V, T),
\]

where

\[
Z(V, T) = \sum_{(i)} e^{-\beta E_i}
\]

and \( \beta = 1/k_B T \), where \( k_B \) is the Boltzmann’s constant and \( E_i \) is the \( i \)-th energy level of the total system. To further elaborate the above relation we need to specify the considered collection of particles first and to introduce a number of approximations (simplifications) to make the approach feasible.

1. Let the molecule be composed of \( k \) atoms and let us consider a system of \( N \) identical and indistinguishable molecules. Molecules are not stiff ‘dead’ bodies but the atoms ‘live’ within the molecules. Thus, to completely specify the location of the molecule in the configuration space, we have to supply then \( 3k \) position coordinates; for the location in the phase space we need to add conjugate momenta. It is convenient to place a coordinate system into the center of mass of the molecule (or any other point called reference point) which means that we have to provide three coordinates of the reference point to describe its translation, and two (in the case of a linear molecule) or three coordinates (for a general molecule) to describe its rotation as a whole. From the total of \( f = 3k \) configurational degrees of freedom of the molecule, there thus remains \( f = 3k - 5 \), resp. \( f = 3k - 6 \), degrees of freedom to get the contribution to the energy from the internal degrees of freedom (vibrations, rotations, etc.).

2. Unless fluids are considered at extreme conditions (i.e., pressures of the order Gbar and higher), it seems reasonable to assume that the internal degrees of freedom of a single molecule (i.e., the motion of the individual atoms within the molecule, and hence the associated energy) are not, on the average, affected by the presence of neighboring molecules (this cannot be, evidently, the case of large flexible molecules). Then, as it is well known
from the basic course of statistical mechanics [11], the canonical ensemble partition function corresponding to the internal degrees of freedom, \( z_{\text{int}} \), can be factored out from the total partition function \( Z \),

\[
Z = \frac{1}{N!} z_{\text{int}}^N \sum_{(i)} e^{-\beta E_i'} \equiv z_{\text{int}}^N Z_N ,
\]

where factor \( 1/N! \) indicates that we are dealing with the collection of indistinguishable particles and the prime says that the total energy does not include the contributions from the internal degrees of freedom.

3. The next question is whether, and then under which conditions, is it acceptable to use the Boltzmann statistics (i.e., classical mechanics) to evaluate the partition function \( Z_N \). Classical physics is a valid approximation to quantum physics whenever the thermal de Broglie wavelength, \( \Lambda = h/|\mathbf{p}| \), where \( h \) is the Planck’s constant and \( \mathbf{p} \) is the momentum of molecules, is much smaller compared to characteristic molecular length scales, e.g., a size of the molecules, \( \sigma \),

\[
\Lambda = \frac{h}{|\mathbf{p}|} \approx \frac{\hbar}{\sqrt{3mk_B T}} \ll \sigma ,
\]

where \( m \) is the mass of the molecule. If \( \sigma \) is of the order of Angströms, then it is easy to verify that the above criterion is satisfied for all systems except, evidently, helium, and then also hydrogen and neon at low temperatures. It is thus justified to use the semi-classical approximation and go over from the summation over the quantum states to integration in the phase space,

\[
Z_N(V, T) = \frac{1}{N! h^{3N}} \int e^{-\beta \mathcal{H}(\{\mathbf{p}^N\},\{\mathbf{q}^N\})} d\{\mathbf{p}^N\} d\{\mathbf{q}^N\} ,
\]

where \( \mathcal{H} \) is the system’s Hamiltonian which is a function of generalized coordinates of molecules, \( \mathbf{q}_i \), and conjugate momenta, \( \mathbf{p}_i \), \( i = 1, \ldots, N \). The shorthand notations \( \{\mathbf{x}^N\} \equiv (\mathbf{x}_1, \mathbf{x}_2, \ldots) \) and \( d\{\mathbf{x}^N\} \equiv d\mathbf{x}_1 d\mathbf{x}_2 \ldots d\mathbf{x}_N \) are used to indicate the collection of variables.

4. The Hamiltonian is given by the sum of the kinetic and potential energies,

\[
\mathcal{H} = K + U ,
\]

where \( K \) depends only on momenta and \( U \) is, with only a few exceptions not encountered in this book, a function of coordinates only,

\[
\mathcal{H} = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + U(\{\mathbf{q}^N\}) ,
\]

for molecules without rotational degrees of freedom; in a more general case the rotational energy of a rigid rotor must be added. The integration over momenta can be performed
immediately yielding

\begin{equation}
Z_N(V, T) = \frac{1}{N!} \left[ \frac{\sqrt{2\pi mk_B T}}{h} \right]^{3N} \int e^{-\beta U(q^N)} d\{q^N\} \equiv \frac{1}{N! \Lambda^3 N} Q_N ,
\end{equation}

where \( \Lambda \) is the above mentioned de Broglie wavelength,

\begin{equation}
\Lambda = \frac{h}{\sqrt{2\pi mk_B T}} ,
\end{equation}

and \( Q_N \) is the so called configurational integral,

\begin{equation}
Q_N = \int e^{-\beta U(q^N)} d\{q^N\} .
\end{equation}

In a general case of nonspherical molecules the rigid rotor contributes to the partition function by

the rotational partition function per molecule the term

\begin{equation}
q_{rot} = \frac{8\pi I K}{h^2} ,
\end{equation}

in the case of linear molecules, and

\begin{equation}
q_{rot} = \left( \frac{8\pi^2 k_B T}{h^2} \right)^{3/2} (\pi I_A I_B I_C)^{1/2}
\end{equation}

in the case of general molecules. Here \( I \) and \( I_K \) are the momenta of inertia.

In the general expression (2.3) for the partition function \( Z \) we have factored out the contribution of
the internal degrees of freedom which implies also the separation of their contribution to the thermodynamic functions. In the limiting case of \( U(q^N) \to 0 \) (ideal gas limit) we get immediately for the configuration integral the result \( Q_N \to V^N \). It is therefore convenient to further factor out also the ideal gas contributions to the thermodynamic properties and focus only on the contributions resulting from the non-zero intermolecular interactions referred to as the excess properties henceforth. Thus, writing the partition function as

\begin{equation}
Z_N(V, T) = Z_N^{id}(V, T) \frac{Q_N(V, T)}{V^N} ,
\end{equation}

where

\begin{equation}
Z_N^{id}(V, T) = \frac{V^N}{N! \Lambda^3 N} ,
\end{equation}

the Helmholtz free energy splits into the corresponding ideal gas and excess terms:

\begin{equation}
\beta A = -\ln Z_N^{id}(V, T) - \ln \frac{Q_N(V, T)}{V^N} \equiv \beta A^{id} + \beta A^{ex} .
\end{equation}

For the ideal gas part we get from (2.15)

\begin{equation}
\frac{\beta A^{id}}{N} = \ln(\rho \Lambda^3) - 1 .
\end{equation}
and hence, using Eqns. (A.8), (A.6), and (A.11), the thermodynamic functions of an ideal gas are obtained:

\[ U_{\text{id}}/N = \frac{3}{2} k_B T, \]  
\[ P_{\text{id}} = \rho k_B T, \]  
\[ \mu_{\text{id}} = k_B T \ln(\rho \Lambda^3), \]

where \( \rho = N/V \) denotes the number density. Finally, by differentiating \( A_{\text{ex}} \) with respect to \( T, V, \) and \( N \), we get expressions for the excess internal energy, pressure, and chemical potential, respectively, by means of the configurational integral:

\[ U_{\text{ex}} = k_B T^2 \frac{1}{Q_N} \left( \frac{\partial Q_N}{\partial T} \right)_{V,N}, \]  
\[ P_{\text{ex}} = k_B T \frac{1}{Q_N} \left( \frac{\partial Q_N}{\partial V} \right)_{T,N} - \rho k_B T, \]  
\[ \mu_{\text{ex}} = - k_B T \ln \frac{Q_N}{V Q_{N-1}}. \]

In the following text, when talking about or discussing thermodynamic functions we will usually have in mind the configurational functions without explicitly emphasizing it unless the cases where it could be misleading; the superscript ‘ex’ will also be therefore omitted.

All the above results can be easily extended to mixtures of fluids consisting of \( N_1 \) molecules of species 1, . . . , and \( N_L \) molecules of species \( L \), \( \sum N_i = N \). The partition function of the mixture is given by

\[ Z_{N,\text{mix}}(V, T, \{N_i\}) = \frac{1}{N_1! N_2! \ldots N_L!} \frac{1}{\Lambda^N} \int e^{-\beta U(\{q_N\})} \prod_{i=1}^{L} d\{q_i^{N_i}\}, \]  

from which all the necessary relations can be derived.

### 2.2 Pairwise additivity

As we have seen in the preceding section, the entire problem of evaluation of the partition function of fluids has been reduced to the evaluation of the \( N \)-particle configuration integral \( Q_N = Q_N(V, T) \). To proceed further one has to specify the potential energy function \( U \) which fully determines the properties of the considered system. Since we consider fluids in equilibrium in the absence of any external field, \( U \) is given only by the energy of mutual intermolecular interactions. In general, one can always write

\[ U = \sum_{i<j} u_{ij} + \sum_{i<j<k} u_{ijk} + \ldots, \]
where the first term gives the energy of all \(i-j\) pairs with \(u_{ij}\) being the pair interaction potential, the second the energy of all triples without the contribution of pairs (indicated by the prime at the summation symbol), etc. It is evident that the more terms in (2.26) are used to estimate the total \(U\), the more complex the problem becomes, defying likely any theoretical treatment. The dilemma to be resolved reads thus as follows:

- to include higher-order potentials to make \(U\) as accurate as possible, which would however vastly complicate the structure of any theory and numerical computations, or
- to resort to the lowest (crudest) approximation of pairwise additivity to make the molecular approach tractable.

There are two strong arguments favoring the latter option: (i) contemporary statistical mechanical theories (including molecular simulations, see Chapter 5) are still not able to efficiently handle systems with triplet and higher-order forces, and (ii) as for the individual contributions to \(U\), quantum chemistry has not been able so far to provide in a closed analytic form, in general, nothing more than (approximate) pair potential models. As a consequence of these two facts, the most common assumption is therefore that of pairwise additivity:

\[
U(\{q^N\}) = \sum_{i<j} u(q_i, q_j) = \sum_{i<j} u(R_i, \Omega_i, R_j, \Omega_j) = \sum_{i<j} u(R_{ij}, \Omega_i, \Omega_j, \Omega), 
\]  

(2.27)

where \(R_{12}\) denotes the separation between the reference points of molecules \(i\) and \(j\), \(R_{ij} = |R_i - R_j|\), and \(\Omega_i\) denotes the orientation of molecule \(i\).

Strictly speaking, this is a rather crude approximation and fails in the case of dense fluids if for \(u(q_i, q_j)\) the true pair potentials (i.e., as determined from the properties of dilute gases) are used. However, if certain effective pair potentials \(u_{\text{eff}}\) are used in (2.27), then this equation represents an acceptable and reasonably accurate approximation to the total configurational energy. Throughout this monograph \(u\) will always mean an effective potential and the subscript ‘eff’ is dropped for simplicity.

Once the effective potential is chosen, the formal road for applying the standard statistical mechanical methods is opened, yet there are still many obstacles impeding a specific computation. One immediate simplification we get if the assumption of pairwise additivity is used in the configuration integral,

\[
Q_N = \int e^{-\beta U(\{q^N\})} d\{q^N\} = \int \ldots \int \exp[- \beta \sum_{i<j} u_{ij}] d\{q^N\} = \int \ldots \int \prod_{i<j} \exp[-\beta u_{ij}] d\{q^N\}. 
\]  

(2.28)

Since \(u(R) \to 0\) for large \(R\) regardless of the actual choice of \(u\), the individual factors in the product go to unity for large \(R\) and this form is thus rather inconvenient for any both mathematical and physical considerations. An elementary trick which makes life much easier is
To replace the exponential function of the pair potential by a function which decays to zero with increasing particle-particle separation. Such a function is the Mayer function $f$, 

$$f(R) = e^{-\beta u(R)} - 1 ,$$  

(2.29)

by means of which the configurational integral is given as

$$Q_N = \int \ldots \int \prod_{i<j} (1 + f_{ij}) d\{q^N\} = \int \ldots \int \left[ 1 + \sum_{i<j} f_{ij} + \sum_{i<j,k<l} f_{ij} f_{kl} + \ldots \right] d\{q^N\} ,$$  

(2.30)

where for simplicity we use the notation $f_{ij} = f(q_i, q_j)$.

2.3 Classes of fluids

To make contact with existing textbooks on liquids and a widespread concept, we will follow in this subsection the traditional path leading to a common classification of fluids [but not necessarily to physically correct models of $u(q_i, q_j)$]!

Accounting for the overall electroneutrality of molecules, physical considerations identify four main types of interactions acting between the molecules of pure fluids:

1. Short-range repulsions that reflect, roughly, the shape and size of molecules (excluded volume effects);
2. Relatively weak and fast decaying (as $1/R^6$ and faster, where $R$ is an intermolecular separation) attractive interactions (called dispersion or van der Waals interactions);
3. Long-range electrostatic interactions (e.g., dipole-dipole) having their origin in the permanent multipoles of molecules;
4. Strong short-range and strongly orientation-dependent attractions identified as hydrogen bonding interactions (H-bonding).

It has thus been common and natural to write an approximate intermolecular interaction model accordingly:

$$u(q_i, q_j) = u_{\text{rep}}(q_i, q_j) + u_{\text{disp}}(q_i, q_j) + u_{\text{multipole-multipole}}(q_i, q_j) + u_{\text{H-bond}}(q_i, q_j)$$

$$\equiv u_{\text{non-el}}(q_i, q_j) + u_{\text{multipole-multipole}}(q_i, q_j) + u_{\text{H-bond}}(q_i, q_j) ,$$  

(2.31)

where the individual terms, the repulsive, $u_{\text{rep}}$, dispersive, $u_{\text{disp}}$, multipolar, $u_{\text{multipole-multipole}}$, and the hydrogen bonding term, $u_{\text{H-bond}}$, will be specified and discussed later.

The intuitive decomposition (2.31) of the total pair interaction $u$ into individual physically meaningful contributions has formed a natural basis for the traditional classification of fluids made up of (relatively) small molecules [34]: general qualitative features of the intermolecular interactions are considered and the fluids are classified in order of increasing complexity of pair interactions:
CHAPTER 2. SYSTEMS OF INTERACTING PARTICLES: BASIC CONSIDERATIONS

1. **Normal fluids.** These fluids are characterized by either negligible or completely missing electrostatic interactions between their molecules,

\[ u_{\text{normal}}(q_1, q_2) = u_{\text{non-el}}(q_1, q_2) . \tag{2.32} \]

Typical examples are, e.g., inert gases and lower hydrocarbons. Two subclasses are often distinguished:

(a) **Simple fluids.** The molecules of these fluids are either spherically symmetric (e.g., argon) or the spherical symmetry of molecules is a very good approximation (as, e.g., methane). Thus,

\[ u(q_1, q_2) = u(R_{12}) . \tag{2.33} \]

(b) **Molecular fluids.** The shape of molecules of these fluids deviates significantly from the spherical symmetry, e.g. lower hydrocarbons, and dependence on the mutual orientation of molecules must be accounted for,

\[ u(q_1, q_2) = u(R_{12}, \Omega_1, \Omega_2) . \tag{2.34} \]

2. **Polar fluids.** Their molecules possess the permanent multipoles which depend on the symmetry of the molecules. The \( u_{\text{multipole-multipole}} \) term is thus added to \( u_{\text{non-el}} \) to account for a significant deviation of the charge distribution from sphericity:

\[ u_{\text{polar}}(q_1, q_2) = u_{\text{non-el}}(q_1, q_2) + u_{\text{multipole-multipole}}(q_1, q_2) . \tag{2.35} \]

To be more specific, we may further distinguish between

(a) **Dipolar fluids** with the dipole-dipole interaction added to \( u_{\text{non-el}} \),

\[ u_{\text{DD}}(q_1, q_2) = \frac{1}{R_{12}^3} \left[ (\mathbf{D}_1 \cdot \mathbf{D}_2) - \frac{3(\mathbf{D}_1 \cdot \mathbf{R}_{12})(\mathbf{D}_2 \cdot \mathbf{R}_{12})}{R_{12}^2} \right] , \tag{2.36} \]

where \( \mathbf{D}_i \) is the dipole moment of molecule \( i \), and

(b) **Quadrupolar fluids**, i.e., the fluids made up molecules whose dipole moment is, due to their symmetry, zero. Only the quadrupole-quadrupole interaction is thus added.

Typical examples of the dipolar fluids are fluids of heterogeneous diatomic molecules; further examples of strong dipolar fluids are, e.g., acetone or acetonitrile. Examples of quadrupolar fluids are homonuclear diatomics or linear triatomics (carbon dioxide).

In a general case when both the dipole and quadrupole moments are non-zero, then also the dipole-quadrupole energy must be added to the interaction energy.

3. **Associating fluids.** At short separations and sharply localized mutual orientations, the attractive interaction between their molecules may become extremely strong (establishing the so called hydrogen bond, H-bond) giving rise to long-living multimers and molecular network. Typical examples are water and alcohols. In addition to being subject to (usually) strong dipole-dipole interactions, their interaction potential includes also thus a certain H-bonding term:

\[ u_{\text{assoc}} = u_{\text{non-el}}(q_1, q_2) + u_{\text{multipole-multipole}}(q_1, q_2) + u_{\text{H-bond}}(q_1, q_2) . \tag{2.37} \]
2.3. CLASSES OF FLUIDS

The above three classes of fluids will be the subject of this introductory textbook, referring to the polar and associating fluids as complex fluids. However, the normal and complex fluids are not the only types of fluids we may encounter. For completeness of the classification one should add two more types: (4) fluids in which the long range charge-charge forces dominate the interactions and (5) fluids for which the internal degrees of freedom of molecules play the decisive role. Into the former group belong metals (e.g., mercury) and molten salts, and into the latter primarily polymer systems and systems of biomolecules referred nowadays to as soft matter.

In the context of the above classification and corresponding interaction potential functions it is necessary to point out that the functions added to the non-electrostatic interactions do not represent any genuine physical forces. In fact, neither multipoles nor the H-bond interaction are real physical entities. They result from mathematical manipulations aiming at a simple description of the electric field outside the set of charges (multipole expansion) or from an attempt to estimate the resulting effect (H-bonding) of the true physical interactions acting between the molecules. As for the latter, no realistic mathematical models have been in fact, for obvious reasons, developed. The $u_{\text{H–bond}}$ term is thus used verbally primarily for an intuitive explanation of certain observed phenomenon of bulk fluids or as an oversimplified caricature of this phenomenon in theoretical models. Nonetheless, the above discussed classification is a useful tool to differentiate between different fluids and understand their properties and differences.

The above classification tends to express (explain) the properties of a more complex fluid in terms of an excess over a less complex (simpler) fluid pointing to a perturbation treatment as a potential and suitable tool for both theory and applications. The observable differences in the behavior of different substances (classes of fluids) seem to reflect differences in the relative strengths of the individual contributions in (2.31) to the total $u$, and the properties of fluids belonging to different classes seem thus to be determined by the different types of predominant interactions. Thus, to be able to understand and, consequently, to predict the macroscopic properties of fluids, it would be necessary to determine the effect of the individual terms contributing to $u$ on the macroscopic behavior. This may seem a natural way to follow and this actually was the path of development of fluid theories till, let us say, the end of the 1980’s.

After developing very successful theories for simple (normal) fluids, problems were encountered to accurately describe the properties of polar fluids. Certain partial useful results were obtained, no question, but, in general, no satisfactory results that would be valid over a wide range of thermodynamic conditions and the range of multipole values (for a review of such theories see, e.g., a monograph by Gubbins and Gray [21]). Furthermore, attempts to attack water, the most prominent example of associating fluids, failed completely. As it will become clear soon, also this failure is very likely due to the entire concept of building up the pair potential in the form (2.31). This concept has been therefore gradually abandoned and has been used only
occasionally, primarily in engineering applications.

2.4 Simple fluid interaction models

So far we have not specified the unquestionable non-electrostatic term in Eqn. (2.31). To be able to proceed further it is time we did it now. Intuitively we may deduce that at short separations the molecules must strongly repel each other to prevent collapse of matter. On the other hand, at large separations they must attract one another to prevent decomposition of matter. This course of \( u(R) \) is schematically shown in Fig. 2.1.

The question is, what is the origin of the attraction and repulsion when the molecules (atoms) are electrically neutral? To answer lies in quantum mechanics because they both are of purely quantal origin. The quantum mechanical perturbation theory shows that the attractive interaction results from the fluctuation of the charge distribution. The actual calculations performed for simple atoms as, e.g. argon, then yield that at large separations the intermolecular interaction assumes the form of the expansion in the inverse even powers of the intermolecular distance,

\[
\begin{align*}
  u(R) \approx - \left( \frac{A}{R^8} + \frac{B}{R^8} + \frac{C}{R^{10}} + \ldots \right) .
\end{align*}
\] (2.36)
In most applications the higher order terms are neglected and only the interaction decaying as \( A/R^6 \), referred to usually as van der Waals (dispersive) interaction, is retained, see Eqn. (2.38) and (2.39).

At close separations the repulsive interaction is easier to understand: In this range the repulsion between two spherically symmetric atoms arises from an overlap of the electron shells. Nonetheless, it is again of purely quantal origin and is associated with the Pauli exclusion principle. There are both quantum mechanical and experimental (scattering of very energetic molecular beams) results available. Experimental results [31] show, in agreement with theoretical estimates [32], that in the narrow range of the short separations \( \ln u(R) \) varies, approximately, linearly with \( R \), i.e.,

\[
\ln u(R) \approx B e^{-kR}.
\]

Unfortunately, no such straightforward (and simple) theoretical results are available for the intermediate region. The total interaction potential we may thus get by stitching the above two results together, see Fig. 2.1,

\[
u(R) = Be^{-kR} - \frac{A}{R^6},
\]

where parameters \( A, B, \) and \( k \) are treated as adjustable. This potential, known under the name Buckingham or exp-6 potential, is usually written in an equivalent form,

\[
u_B(R) = \frac{\epsilon}{1 - 6/\alpha} \left[ 6 \frac{\epsilon^\alpha(1-R/R_m)}{\alpha} - \left( \frac{R_m}{R} \right)^6 \right],
\]

to give to the parameters a physical meaning: \( R_m \) is the potential minimum at which the energy equals \( -\epsilon \), and \( \alpha \) determines the steepness of the repulsive part of the potential. For real simple fluids \( \alpha \) varies, roughly, between 6 and 15; the larger \( \alpha \), the the steeper potential. The potential exhibits an unphysical maximum at a very close separation \( R_{\text{max}} \) where the repulsion is supposed to be extremely strong. The potential is therefore usually modified by placing a hard core to \( r_{\text{max}} \). The exp-6 potential with this hard core is referred to in the literature as a modified Buckingham potential. Neither this modification seems very appropriate because the repulsion should become stronger with decreasing \( R \) and not weaker. Computer experiments show that even at very high temperatures the physically accessible separations are larger not only than \( R_{\text{max}} \) but also than the inflection point \( R_{\text{inf}} \). Thus, unless extremely high temperatures are considered, a unique physically plausible modification of the original potential (2.39) is

\[
u_{\text{exp6}}(R) = \begin{cases} +\infty & \text{for } R < R_{\text{inf}} \\ \frac{\epsilon}{1 - 6/\alpha} \left[ 6 \frac{\epsilon^\alpha(1-R/R_m)}{\alpha} - \left( \frac{R_m}{R} \right)^6 \right] & \text{for } R > R_{\text{inf}} \end{cases}
\]

with the negative second derivative down to \( R_{\text{inf}} \). This modification is used later in this textbook and referred to, for simplicity, also as an exp-6 potential.
Figure 2.2: The course of the Lennard-Jones potential.

The exponential form of the repulsive term is rather cumbersome to handle in analytic calculations. Since the effect of the repulsive interaction applies only to rather a very narrow range of separation, see Fig. 2.37; Eqn. (2.37) can be further simplified. Performing a simple mathematical manipulation we then get that, over a narrow range of separations, Eqn. (2.37) can be cast to a simpler approximate form,

$$u(R) \approx BR^m - m.$$  \hfill (2.41)

This approximate result has served as the basis for developing various fluid models, all of them being motivated by deriving simpler, mathematically more easily tractable functions. The most popular has become the $m - 6$ potential,

$$u(R) = \frac{B}{R^m} - \frac{A}{R^6}.$$  \hfill (2.42)

This potential, with $m$ properly adjusted, performs surprisingly well for simple fluids as neon, argon, or methane [7]. A prominent position among the $m - 6$ potentials is occupied by the choice $m = 12$, known as the Lennard-Jones (LJ) potential. This potential, which can be rearranged into the form

$$u_{\text{LJ}}(R) = 4\epsilon \left[ \left( \frac{\sigma}{R} \right)^{12} - \left( \frac{\sigma}{R} \right)^6 \right],$$  \hfill (2.43)

has gained, due to the $m - n$ combination which simplifies many calculations, enormous popularity and has been used as a prototype of non-electrostatic interactions. But again, it is necessary
2.4. SIMPLE FLUID INTERACTION MODELS

Figure 2.3: Simplifications of the Lennard-Jones potential: square-well model (a) and the hard-sphere model (b).

To remind that there is no theoretical reason justifying this choice, it is simply only matter of convenience. The course of this potential is shown in Fig. 2.2 from which the interpretation of the parameters $\epsilon$ and $\sigma$ becomes also evident: $\epsilon$ is the depth of the potential well and $u(\sigma) = 0$ with $\sigma$ being often interpreted as a collision diameter.

Somewhat outside the main stream of the above models falls another rather useful realistic potential model, the hard-core Yukawa (Y) potential,

\[
\begin{align*}
\text{if } R < \sigma & & u_Y(R) = +\infty \\
\text{if } R > \sigma & & u_Y(R) = -\frac{\sigma}{R} e^{\sigma(1-R/\sigma)}
\end{align*}
\]

originally arising from elementary particle problems. Unlike the two-parameter LJ potential it contains an additional parameter, $z$, which controls the range of the potential and makes thus the potential very flexible. For $z \to +\infty$ it approaches the hard sphere potential (see Eqn. (2.46)) whereas for $z \to 0$ it approaches the Coulomb potential and is therefore also referred to as a screened Coulomb potential. An additional appealing feature of the Y potential is its ability, by a judicial choice of its parameters, to approximate fairly well common variety of realistic simple potential models.

Despite relative simplicity of the above potential models, their continuous course makes analytic calculations very complex or even impossible. Further simplifications, this time based purely on physical arguments, has therefore been made, see Fig. 2.3, the repulsive part of $u(R)$ is very steep and may be thus approximated by a vertical line defining a hard core of the molecule. Outside this core there must be an attractive region, a well; the simplest choice is a square-well (SW). Finally, a very week attraction at large separations, i.e., outside this well,
may be neglected so that we may write
\[ u_{SW}(R) = \begin{cases} +\infty & \text{for } R < \sigma \\ -\epsilon & \text{for } \sigma < R < \lambda\sigma \\ 0 & \text{for } R > \lambda\sigma \end{cases} \]
where \( \lambda \) defines the range of the potential. One may argue that the interaction energy should decay with increasing separation and that the approximation \( u(R) = \text{const} \) for \( \sigma < R < \lambda\sigma \) is too crude. However, other choices as, e.g., the triangular well, make all calculations more complex again and there is therefore no reason for their use in theoretical considerations.

To summarize, when dealing with simple real fluids, the LJ or exp-6 potentials are used, whereas in theoretical considerations the SW potential is commonly used for its simplicity and yet its sound physical footing. For reasons which will be given later, the Y model has also begun attracting attention recently both in theoretical computations and applications.

Perhaps surprisingly, we have not mentioned yet another simple and very important model, the hard sphere (HS) potential,
\[ u_{HS}(R) = \begin{cases} +\infty & \text{for } R < \sigma \\ 0 & \text{for } R > \sigma \end{cases} \]
In fact, when introducing simple models of fluids in textbooks and monographs then, nearly as a rule, the HS model is always introduced at the very beginning and the same applies to general various non-spherical hard body models which represent the simplest purely repulsive interaction models. We intentionally have not done so till this moment because it cannot be clear at first sight what purpose such a model, without any attractive part, may be good for. Nonetheless, the hard body models have played an extremely important role in theories of fluids and their applications but why it will become clear later when they emerge quite naturally within general perturbation theory concepts.

### 2.5 Realistic potential models

With the advance of computer technology, quantum chemical computation methods have also made considerable progress in the development of reasonably accurate effective pair potentials, however, in a form which completely differs from that of equation (2.31).

Small (neutral) atoms may be viewed as spherical bodies with interactions depending only on their relative distance. For their mutual interaction model some of the \( u_{\text{non-el}} \) models listed above may be chosen. When such atoms form molecules then, roughly speaking, their electron shell gets evidently distorted and within the molecules we can localize the nuclei of the atoms and negative charge densities. Consequently, molecules are pictured as bodies made up of individual
2.5. REALISTIC POTENTIAL MODELS

Figure 2.4: Schematic representation of molecules made up of interaction sites.

atoms or groups of atoms or, in general, certain interaction sites and, regardless of the observed macroscopic properties and their possible peculiarities, the potential functions are then written in a uniform way as a sum of interactions between these sites (site-site model), see Fig. 2.4,

\[
u(q_1, q_2) = \sum_{i,j} \mu_{site1,i-site2,j}(|r^{(i)}_1 - r^{(j)}_2|) \equiv \sum_{i,j} \mu_{ij}(r_{ij}), \quad (2.47)
\]

where \(\mu_{site,i-site,j}\) is a spherically symmetric (!) simple interaction acting between site \(i\) on molecule 1 and site \(j\) on molecule 2 with \(r^{(i)}_k\) being their position vectors. It is also usually assumed, for the sake of simplicity, that for relatively small molecules the geometrical arrangement of the sites is fixed within the molecules (rigid monomer).

It remains to specify the site-site interactions. From that said above it follows that the interaction sites are the seat of two types of interactions: (1) non-electrostatic interaction generating a strong repulsion at short separations and a weak (van der Waals) attraction at medium separations, \(\mu_{non-el}\), and (2) the long-range Coulombic interaction, \(\mu_{Coul}\), generated by an effective charge \(q\). A common realistic pair potential assumes thus the following explicit form

\[
u(q_1, q_2) = \nu(R_{12}, \Omega_1, \Omega_2) = \nu_{non-el}(q_1, q_2) + \nu_{Coul}(q_1, q_2)
\]

\[
= \sum_{i \in \{1\}} \sum_{j \in \{2\}} \left\{ \nu_{non-el}(|r^{(i)}_1 - r^{(j)}_2|) + q^{(i)}_1 q^{(j)}_2 \frac{|r^{(i)}_1 - r^{(j)}_2|}{|r^{(i)}_1 - r^{(j)}_2|} \right\}, \quad (2.48)
\]

where \(q^{(i)}_k\) is the partial charge of site \(i\) of molecule \(k\).

Potentials for different compounds thus differ only in the geometrical arrangement of the interaction sites and in the strengths of the individual site-site interactions; effects which might
be identified as, e.g., the dipole-dipole interaction or hydrogen bonding are included implicitly and result as a net effect of the complex intermolecular force field. To find a link between the general model (2.47) and the intuitive physical model (2.31), it is sufficient to realize that multipole moments are not real physical entities but that they arise only from mathematical manipulations (expansion in inverse powers of the charge-charge separation) attempting to approximate the force field of a spatially distributed set of charges; the same applies also to the H-bond interaction. While being reasonable at large separations [and numerically equivalent to the form of (2.47)], it is important to bear in mind that at short intermolecular separations this approximation is inadequate unless a large number of multipole-multipole interactions terms are incorporated.

To demonstrate the above frequently ignored difference between the field created by a set of charges and that due to an idealized point dipole, let us consider a simple example: two charges \( \pm e/4 \), where \( e \) is the elementary charge, a distance \( l = 2 \) Å apart, see Fig. 2.5. The dipole moment of this pair of charges, \( D = ql \), is about 2.4 Debay \((1 \text{ Debay} = 3.33564 	imes 10^{-30} \text{ Cm})\) which roughly corresponds to the dipole moment of acetone (cf., water has \( D = 1.8 \text{ Debay} \), whereas the strong dipolar compound acetonitrile has \( D \) about 5 Debay).

\[ u_{qq}(\mathbf{R}) = q \left( \frac{1}{r_+} - \frac{1}{r_-} \right). \]  

(2.49)
2.5. REALISTIC POTENTIAL MODELS

For $R \gg l$ we have $r_+ \approx R$ and also $r_- \approx R$ so that we may thus write

$$u \approx \frac{qL \cos \theta}{R^2} = \frac{D \cos \theta}{R^2} \equiv u_{\text{dip}}(R). \quad (2.50)$$

Figure 2.6: Comparison of the field of the set of two charges (filled symbols) with that created by their dipole moment (open symbols) at $P$, Fig. 2.5, for a set of distances in dependence on the orientation.

In Fig. 2.6 we compare the field $u_{eq}$ with the that of an ideal dipole as a function of $\theta$ for a series of distances $r$. As it is discernible, at close separations the two fields are really considerably different; consequently, also the interaction energy of two such molecules at short separations (first coordination shell) must be different. The essential point is that the deficiency in an estimation of the field created by a set of charges by that of the point multipoles would have no substantial implications and could be ignored if the long-range interactions did play the dominant role for polar and associating fluids. However, and this must be borne in mind, this does not seem to be the case and the above facts and the existing data seem to strongly oppose this possibility: As it will be shown later in Chap. 6., it is the local force field around the molecules what primarily determines the properties of dense fluids and not the long-range part of interactions.
Chapter 3

Systems of interacting particles: Correlation functions

In common laboratory experiments we characterize fluids by their thermodynamic functions: pressure, density, etc. At the level of molecules the system is fully described in the configuration space by the positions (and orientations in the case of nonspherical molecules) of all the molecules from which the configuration integral can be evaluated. It would be evidently foolish to keep track of the time development of positions of all particles and operate with this redundant and, in principle, inaccessible information. Another and simpler, but sufficient formalism must be used. The quantity(ies) to be used must satisfy two criteria: (i) it/they must be accessible by an experiment, and (ii) it/they must contain enough information to make it possible to evaluate all desired (and potentially measurable) quantities. These criteria are satisfied by correlation functions that provide information on the probability of the mutual distribution of the molecules in the system.

Before using the formal mathematical language to define correlation functions, let us use first simple physical reasoning.

We will consider a system of particles interacting via a simple potential $u(R_{12})$. Let us pick up an arbitrary particle and ask the question: How many particles can be find anywhere a distance $r$ apart (more precisely, in a volume element $(r, r + \Delta r)$ from the chosen particle? If the molecules did not affect one another (did not interact), the number of particles would be the number density times the volume element, i.e.,

$$n(r) = \frac{N - 1}{V} 4\pi r^2 \Delta r .$$

(3.1)

Now, in the general case of non-zero interactions, we will account for the effect of the interactions by introducing a function $g(r)$:

$$n(r) = \frac{N - 1}{V} 4\pi g(r)r^2 \Delta r .$$

(3.2)
3.1 Molecular correlation functions

What the function \( g(r) \) may be good for? Let us try to determine, for instance, the total internal energy. The energy of the chosen particle with all the particles in the shell \((r, r + \Delta r)\) is \( u(r)n(r) \). From Eqn. (3.2) we then get for the energy of the chosen particle with all other particles of the system

\[
U_1 = \frac{N - 1}{N} 4\pi \int_0^{+\infty} g(r)u(r)r^2 \, dr .
\] (3.3)

To get the energy of the entire system we have to sum up over all particles of the system, i.e.,

\[
U = \frac{N(N - 1)}{2N} 4\pi \int_0^{+\infty} g(r)u(r)r^2 \, dr ,
\] (3.4)

and thus

\[
U/N = 2\pi \rho \int_0^{+\infty} g(r)u(r)r^2 \, dr .
\] (3.5)

In other words, we see that if we know the function \( g(r) \), called the correlation function, then we can simply determine the internal energy. This is one of the features of the correlation functions introduced rigorously below.

### 3.1 Molecular correlation functions

By definition, the quantity

\[
\frac{1}{Q_N} e^{-\beta U(N)} \, dq_1 \cdots dq_N ,
\] (3.6)

provides the probability of finding particle 1 in the volume element \( dq_1 \) around \( q_1 \), particle 2 in the volume element \( dq_2 \) around \( q_2 \), \ldots, and particle \( N \) in the element \( dq_N \) around \( q_N \). Hardly may we be interested in this complete information of the arrangement of all particles. More reasonable function is the information on the arrangement of a small group of particles only, let us say particles 1 to \( k \) around points \( q_1, \ldots, q_k \) regardless of the positions of the remaining \( N - k \) particles. Such information we get from (3.6) by integrating over the positions of the uninterested particles \( k + 1, \ldots, N \),

\[
P^{(k)}(\{q^k\}) = \frac{1}{Q_N} \int e^{-\beta U(\{q^N\})} \, dq_{k+1} \cdots dq_N .
\] (3.7)

Since we study the system of indistinguishable particles, it is immaterial whether there is the particle with a specific label around the point \( q \); what we need to know is that there is a particle around this point. We must therefore multiply \( P^{(k)} \) by the appropriate combinatorial factor to get the distribution function of the order \( k \),

\[
\rho^{(k)}(q_1, \ldots, q_k) = \frac{N!}{(N - k)!} \frac{1}{Q_N} \int e^{-\beta U(\{q^N\})} \, dq_{k+1} \cdots dq_N ,
\] (3.8)

which is normalized as

\[
\int \rho^{(k)}(q_1, \ldots, q_k) \, dq_1 \cdots dq_k = \frac{N!}{(N - k)!} .
\] (3.9)
When the intermolecular interaction is small, i.e., $U \to 0$, then the integration over the positions of particles $k+1, \ldots, N$ contributes the factor $V^{N-k}$ to $\rho^{(k)}(q_1, \ldots, q_v)$ so that

$$\rho^{(k)}(q_1, \ldots, q_k) \approx V^{N-k} \rho(q_1, \ldots, q_v), \quad k \ll N.$$  

It is therefore convenient to factor out this limiting value from $\rho^{(k)}(q_1, \ldots, q_k)$ and introduce the so-called correlation functions $g^{(k)}$,

$$\rho^{(k)}(q_1, \ldots, q_k) = \rho^{k} g^{(k)}(q_1, \ldots, q_k).$$  

These functions do describe the correlation between the particles in the system: When the intermolecular interaction tends to zero then there is no correlation between the positions of particles and the correlation functions tend to unity. For $k = 1$ the distribution function $\rho^{(1)}$ equals $\rho$ as it should be because in the absence of an external field the distribution of particles in fluids is uniform, i.e., $g^{(1)} = 1$.

What the correlation functions of different order may be good for? Let us consider, for simplicity, a simple liquid, i.e., $q \equiv R$, and make use of the assumption of pairwise additivity in (2.27):

$$k_B T^2 \left( \frac{\partial \ln Q_N}{\partial T} \right)_{V,N} = \frac{1}{Q_N} \int e^{-\beta U(R)} \left( \sum_{i<j} u(R_{ij}) \right) d\{R^N\}$$

$$= \frac{N(N-1)}{2Q_N} \int e^{-\beta U(R)} u(R_{12}) d\{R^N\}$$

$$= \frac{1}{2} \int \int u(R_{12}) \rho^{(2)}(R_1, R_2) dR_1 dR_2$$

$$= 2\pi N \rho \int_0^{+\infty} u(R) g^{(2)}(R) R^2 dR.$$  

For the configurational energy per particle we thus get

$$U/N = 2\pi \rho \int_0^{+\infty} u(R) g^{(2)}(R) R^2 dR,$$  

i.e., the same result which we derived intuitively at the beginning of this Chapter, $g \equiv g^{(2)}$. Thus, as we see, to determine the excess internal energy in such a system it is sufficient to know the correlation function of the second order. But what about pressure? It is given by the derivative of the configuration integral with respect to volume. However, this is not so straightforward; first, it is necessary to evaluate the derivative with respect to $V$ which can be worked out by employing the transformation

$$R = R' V^{(1/3)},$$

[1] There is no unique usage of the terms distribution and correlation functions; some authors refer to these correlation functions as distribution functions.
3.1. MOLECULAR CORRELATION FUNCTIONS

which enables us to express explicitly the dependence of $Q_N$ on $V$,

$$Q_N = V^N \int_0^1 \ldots \int_0^1 e^{-\beta U(R^N)} d\{R^N\}.$$  \hspace{1cm} (3.14)

Then,

$$P = \frac{Nk_BT}{V} - \frac{V^2}{2} \int_0^1 \ldots \int_0^1 \frac{du(R_{12})}{dV} \rho^{(2)}(R_1, R_2) dR_1 dR_2.$$  \hspace{1cm} (3.15)

Since

$$\frac{du(R_{12})}{dV} = \frac{R_{12}}{3V} \frac{du}{dR_{12}},$$

we finally get

$$\frac{\beta P}{\rho} = 1 - \frac{2}{3} \pi \beta \rho \int_0^{+\infty} \frac{du}{dR} g(R) R^2 dR.$$  \hspace{1cm} (3.16)

This equation is usually referred to as the virial equation of state because it can be derived also from the virial theorem of classical mechanics [35].

Both the internal energy and pressure are mechanical properties. To complete this exercise, we do the same for the chemical potential to obtain an entropic quantity. To be able to evaluate the fraction $Q_N/Q_{N-1}$, see Eqn. (2.24), let us split the total interaction energy into the energy of a set of $(N-1)$ molecules and the energy of the remaining single molecule, labeled as $N$, with the remaining $(N-1)$ molecules:

$$U(\{q^N\}) = U(\{q^{N-1}\}) + \Psi(q_N).$$  \hspace{1cm} (3.17)

Then we can write

$$e^{-\beta \mu_{ex}} = \frac{1}{VQ_{N-1}} \int e^{-\beta U(\{q^{N-1}\})} e^{-\beta \Psi(q_N)} d\{q^N\}$$

$$= \frac{N(N-1)}{V} \int g(q_1, q_N) e^{-\beta \Psi(a_N)} dq_1 dq_N.$$  \hspace{1cm} (3.18)

This expression is particularly convenient for determination of the chemical potential in molecular simulations, see Chap.4, but not for theoretical considerations. Instead of considering one particle separately as above, we can introduce a scaling parameter $\xi$ which gradually couples such a selected particle into the system,

$$u(R, \xi) = \xi u(R).$$  \hspace{1cm} (3.19)

Then we can derive another expression for $\mu$ [5]:

$$\beta \mu = \ln(\rho \Lambda^3) + 4\pi \beta \rho \int_0^1 d\xi \int_0^{+\infty} \frac{\partial u(R, \xi)}{\partial \xi} g(R, \xi) R^2 dR.$$  \hspace{1cm} (3.20)
In a general case of fluids made up of rigid molecules with the interaction \( u(q_1, q_2) \) then the above considered thermodynamic functions are given by [5]

\[
\frac{\beta P}{\rho} = 1 - \frac{2}{3} \pi \beta \int g(q_1, q_2) R_{12} \frac{\partial u(q_1, q_2)}{\partial R_{12}} \, dq_1 \, dq_2
\]

(3.21)

\[
E = \frac{\rho^2}{2} \int u(q_1, q_2) g(q_1, q_2) \, dq_1 \, dq_2
\]

(3.22)

\[
\mu = k_B T \ln \Lambda \int_a^b d\xi \int \frac{\partial u(\xi; q_1, q_2)}{\partial \xi} g(q_1, q_2) \, dq_1 \, dq_2
\]

(3.23)

To summarize, the pair correlation function contains the complete information on the systems with pairwise additive interaction; in addition to providing information on the structure of the system, it makes it possible to determine also all its thermodynamic properties. Similarly, if we deal with a system with non-additive triplet interaction then, in addition to the pair correlation function, we need also the triplet correlation function \( g^{(3)} \) to obtain the complete description of the system, and so on.

So far we have used the canonical ensemble formalism. In some cases it may be however more advantageous to operate within the grandcanonical formalism. In this case the probability of finding \( N \) particles in the system at a given \( T, V \), and \( \mu \) is given by

\[
P_N = \frac{1}{\Xi} e^{\beta N \mu} Z_N(V, T) = \frac{1}{\Xi} \frac{\phi^N}{N!} Q_N(V, T) ,
\]

(3.24)

where

\[
\phi = \Lambda^{-3} e^{\beta \mu} ,
\]

(3.25)

and the grandcanonical partition function is given by

\[
\Xi = \sum_{N \geq 0} \frac{\phi^N}{N!} Q_N(V, T) .
\]

(3.26)

Distribution functions are then introduced in a way similar to that in the canonical ensemble. The \( k \)-particle distribution function is defined as

\[
\rho^{(k)} = \frac{1}{\Xi} \sum_{N \geq k} \frac{\phi^N}{(N-k)!} \int e^{-\beta U(q)} \, dq^{N-k} ,
\]

(3.27)

and normalized as [c.f. (3.9)]

\[
\int \rho^{(k)}(q_1, \ldots, q_k) \, dq_1 \ldots dq_k = \frac{N!}{(N-k)!} .
\]

(3.28)

In the limiting case of \( U_N \rightarrow 0 \) the behavior of \( \rho^{(k)} \) is the same as in the canonical ensemble with the only difference that now instead of a fixed number of particles there stands the ensemble average, \( N \rightarrow \langle N \rangle \). Thus,

\[
\rho^{(k)} \xrightarrow{U_N \rightarrow 0} \left( \frac{\langle N \rangle}{V} \right)^h \equiv \langle \rho \rangle^k ,
\]

(3.29)
which leads to the following definition of the correlation functions:

$$
\rho^{(k)}(q_1, \ldots, q_k) = \langle \rho \rangle^k g^{(k)}(q_1, \ldots, q_k) .
$$  (3.30)

Whereas the virial EOS requires, for its derivation, the assumption of pairwise additivity, the grandcanonical ensemble makes it possible to derive another, and more general form of EOS without resorting to this assumption. It is based on the relation between the fluctuations in the number of particles and the isothermal compressibility. From the definition of the average number of particles,

$$
\langle N \rangle = \sum N \mathcal{P} ,
$$  (3.31)

we get, after inserting for $\mathcal{P}$ from (3.26) and making use of the relation

$$
\langle N \rangle = \frac{1}{\beta} \left( \frac{\partial \ln \Xi}{\partial \mu} \right)_{T,V} ,
$$  (3.32)

the result

$$
\left( \frac{\partial \langle N \rangle}{\partial \mu} \right)_{T,V} = \beta \left[ \langle N^2 \rangle - \langle N \rangle^2 \right] \equiv \beta \langle (\Delta N)^2 \rangle ,
$$  (3.33)

where $\Delta N = N - \langle N \rangle$ denotes the fluctuation. On the other side, from the normalizing condition (3.28) we immediately get

$$
\int \int [\rho^{(2)}(1, 2) - \rho^{(1)}(1)\rho^{(1)}(2)]d(1)d(2) = < (\Delta N)^2 > - < N > .
$$  (3.34)

Putting now together this result and equation (3.33) and making use of the thermodynamic relation

$$
\left( \frac{\partial P}{\partial \rho} \right)_{T,V} = \langle N \rangle \left( \frac{\partial \mu}{\partial \langle N \rangle} \right)_{T,V}
$$  (3.35)

we finally get

$$
k_B T \left( \frac{\partial \rho}{\partial P} \right)_{T,V} = 1 + \rho \int [g(q_{12}) - 1] dq_{12} .
$$  (3.36)

This equation is called the the compressibility equation of state.

Above we have introduced the pair correlation function using two different approaches: an intuitive and formal statistical-mechanical. There is a third way which is useful in a number of theoretical considerations and which is based on the relation between the $\delta$ functions and particle’s density. Strictly speaking, when introducing the correlation functions in this way we should distinguish between continuous systems (using the Dirac $\delta_D$ function) where the term ‘density’ refers to the number of particles in a volume element, and discrete systems (using $\delta$ function with values 0 and 1) where we can locate particles “at the point”. However, the final results are identical and for the sake of simplicity we will not make any difference between these two cases.
In a system of $N$ particles enclosed in a volume $V$ the instantaneous number of particles in point $\mathbf{q}$ is

$$N(\mathbf{q}) = \sum_{i=1}^{N} \delta^D(\mathbf{q}, \mathbf{q}_i).$$  \hfill (3.37)

Ensemble average of this quantity is

$$\langle N(\mathbf{q}) \rangle = \frac{1}{Q_N} \int e^{-\beta U(\{\mathbf{q}_i\})} \delta^D(\mathbf{q}_1, \mathbf{q}_i) \, d\{\mathbf{q}_1\} = \frac{1}{Q_N} \int e^{-\beta U(\{\mathbf{q}_i\})} \, d\mathbf{q}_2 \ldots \mathbf{q}_N,$$  \hfill (3.38)

so that, by comparing this result with definition (3.8), $\langle N(\mathbf{q}) \rangle$ equals one-particle distribution function $\rho^{(1)}$.

In a similar way we may define the conditional probability of finding particle 2 at $\mathbf{q}_2$ provided that there is particle 1 at $\mathbf{q}_1$:

$$\langle N(\mathbf{q}_1) N(\mathbf{q}_2) \rangle = \sum_{i=1}^{N} \sum_{j=1}^{N} \delta^D(\mathbf{q}_1, \mathbf{q}_i) \delta^D(\mathbf{q}_2, \mathbf{q}_j),$$  \hfill (3.39)

where the prime indicates that the terms with $i = j$ must be omitted and which can easily be shown to be equal to $\rho^{(2)}$. Another useful function which can be defined using this formalism is the correlation between fluctuations of density,

$$G(\mathbf{q}_1, \mathbf{q}_2) = \langle \Delta N(\mathbf{q}_1) \Delta N(\mathbf{q}_2) \rangle = \langle N(\mathbf{q}_1) N(\mathbf{q}_2) \rangle - \langle N(\mathbf{q}_1) \rangle \langle N(\mathbf{q}_2) \rangle,$$  \hfill (3.40)

referred to also as the density-density correlation function. For homogeneous systems we then get

$$G(\mathbf{q}_1, \mathbf{q}_2) = \langle N(\mathbf{q}_1) N(\mathbf{q}_2) \rangle - \rho^2 = \rho^2 [g(\mathbf{q}_1, \mathbf{q}_2) - 1] = \rho^2 h(\mathbf{q}_1, \mathbf{q}_2),$$  \hfill (3.41)

where $h$ is the so-called total correlation function playing the important role in theories of fluids, see Chapter 5.

Above we have introduced the pair correlation function but do not know yet how to obtain it and how it may look out. Nonetheless, simple physical arguments make it possible to draw its certain general features. In the limiting case of an ideal gas there is no correlation between

![Figure 3.1: Schematic representation of the correlation functions of the (a) gas, (b) liquid, and (c) solid (along the lattice edge) phases.](image)
molecules and therefore \( g(r) = 1 \). On the other side of the density range, the molecules are trapped in the lattice points. Examining then the crystal along one of the crystal axis, the probability of finding a molecule is zero unless we hit the lattice point; at the lattice points we get thus \( \delta \)-function like behavior. Both these limiting correlation functions are shown in Fig. 3.1. Another low density limiting case we can get if in definition (3.8) for \( k = 2 \) we factor out from the total interaction energy the direct interaction between particles 1 and 2:

\[
g(q_1, q_2) = e^{-\beta u(q_1, q_2)} \int e^{-\beta U(q)} \frac{dN}{q_3 \ldots q_N} = e^{-\beta u(q_1, q_2)} y(q_1, q_2),
\]

where \( y \) is the so called background correlation function. In the lowest order density expansion of \( y \) gives \( y \rightarrow 1 \) so that at finite low densities (real gas) the Boltzmann factor provides an estimate of the pair correlation function:

\[
g(q_1, q_2) \cong e^{-\beta u(q_1, q_2)}; \quad \rho \approx 0.
\]

Equation (3.43) relates directly the pair correlation function to the pair potential at low densities. This low density relation motivates, in analogy with (3.43), a certain generalization for all densities,

\[
g(q_1, q_2) = e^{-\beta w^{(2)}(q_1, q_2)} ,
\]

where the function \( w^{(2)}(q_1, q_2) \) plays the role of the potential and which can be further split into the direct, \( u(q_1, q_2) \), and indirect, \( W(q_1, q_2) \), parts,

\[
w^{(2)}(q_1, q_2) = u(q_1, q_2) + W(q_1, q_2) .
\]

Function \( W(q_1, q_2) \) is called the potential of the mean force (PMF) and is one of the most important functions in theories of electrolytes.

A typical course of \( g(R) \) for a simple liquid represented by the LJ potential is shown in Fig. 3.2. As we see, the first peak is located, approximately, at the minimum of the potential. Further, the height of the peaks are decreasing, indicating a gradual loss of the configurational arrangement. In other words, liquids exhibit only the short-range ordering which is practically lost beyond, approximately, 6\( \AA \) in this case. In general this holds true for distances of about 5\( \times \) the size of the molecule.

The location of the first minimum, \( R_{\text{min},1} \), defines the region of the first coordination shell and the integration of \( g(R) \) up to \( R_{\text{min},1} \) gives, according to the definition, the number of molecules in the first shell. The running coordination number, \( N_R \), is defined as

\[
N_R = \rho \int_0^R g(R) \, dR .
\]

For simple liquids this is about 12 at normal conditions.

\footnote{Some authors use the term PMF for the total interaction \( w \).}
CHAPTER 3. SYSTEMS OF INTERACTING PARTICLES: CORRELATION FUNCTIONS

3.2 Partial correlation functions

The molecular pair correlation function \( g(q_1, q_2) \) is, in general, function of six variables. Unless an analytic solution of whatever method is obtained, and this is not the case, such a function is quite hard only to present. For instance, it can be evaluated, in principle, from molecular simulations but only to provide tables of a function of six variables is practically impossible.

One possibility, used extensively in the 1970’s and 1980’s, is to expand \( g(q_1, q_2) \) in the spherical harmonics [21], see Chap. 5. The target are then the expansion coefficients which are functions of only one variable, distance. However, this expansion converges relatively slow, particularly for discontinuous potentials, and may find application only in specific theoretical methods.

The most common way to characterize the structure of fluids using a reduced number of variables is by means of certain partial correlation functions. It means, the functions for a fixed partial set of variables:

\[
g(x^k_1, x^k_2) = \int _{\{x^k_1, x^k_2\} = \text{const}} g(q_1, q_2) \, dq_1 \, dq_2.
\]  

(3.47)

An early example of such functions are radial slices through \( g(q_1, q_2) \), see e.g., [36]. With respect to the site-site form of the general Hamiltonian it seems reasonable, from the physical point of view, to know the probabilities of finding of the mutual separation of the individual site with
3.2. PARTIAL CORRELATION FUNCTIONS

The most common partial correlation functions are thus the site-site correlation functions defined as

$$g_{ij}(r) = \int_{|r_i^1 - r_j^2| = \text{const}} g(q_1, q_2) \, dq_1 \, dq_2,$$

where $|r_i^1 - r_j^2| = r$ is the separation between site $i$ on molecule 1 and site $j$ on molecule 2.

To characterize the structure of fluids in terms of the site-site correlation function makes sense if the interaction potential model is also of the site-site type. In this case one can immediately, without necessity of any formal derivation, write down the result for the internal energy,

$$\frac{E}{N} = 2\pi \rho \sum_{i=1}^{L} \int_{0}^{+\infty} u_{ij}(r_{ij}) g_{ij}(r_{ij}) r_{ij}^2 \, dr_{ij},$$

where $L$ is the member of sites of the molecule. However, such a simple result cannot be thus obtained for pressure. It is evident that by integration of $g(q_1, q_2)$ over $dq_1 \, dq_2$ at whatever constrain some pieces of information are lost and the knowledge of the site-site correlation function only is not sufficient and must be supplemented by additional information on the structure. Such an expression for pressure in terms of the site-site functions $g_{ss}$ was first derived by Nezbeda [37] for homonuclear dumbells and later on extended for any site-site interaction.
By comparing (3.16) and (3.22) we see that the crucial quantity to be evaluated is
\[
R_{12} \frac{\partial u(q_1, q_2)}{\partial R_{12}} = \sum_{\alpha, \beta} R_{12} \frac{\partial \alpha \beta}{\partial R_{12}} \frac{du^{\alpha \beta}(r_{12}^{\alpha \beta})}{d(r_{12}^{\alpha \beta})},
\]
where we recall that all differentiations are performed at fixed orientations. From the relation (see Fig. 3.3)
\[
r_{12}^{\alpha \beta} = R_{12} + L_1^\alpha - L_2^\beta,
\]
it is easy to find that
\[
R_{12} \frac{\partial r_{12}^{\alpha \beta}}{\partial R_{12}} = \frac{R_{12} \cdot r_{12}^{\alpha \beta}}{r_{12}^{\alpha \beta}} \equiv (R_{12} \cdot \nu^{\alpha \beta}).
\]
Inserting now (3.50) and (3.52) into (3.23) we get:
\[
\frac{\beta P}{\rho} = 1 - \frac{\beta \rho}{6V} \sum_{\alpha, \beta} \int g(q_1, q_2) \frac{du^{\alpha \beta}}{dr_{12}^{\alpha \beta}} (R_{12} \cdot \nu^{\alpha \beta}) dq_1 dq_2.
\]
In each integral for a given pair \(\alpha\) and \(\beta\) we perform now a coordinate transformation such that the reference points will coincide with the sites \(\alpha\) and \(\beta\) (for details see [37]). Consequently,
\[
\frac{\beta P}{\rho} = 1 - \frac{2\pi \beta \rho}{3} \sum_{\alpha, \beta} \int g(q_1, q_2) \frac{du^{\alpha \beta}}{dr_{12}^{\alpha \beta}} (R_{12} \cdot \nu^{\alpha \beta})(r^{\alpha \beta})^2 d(\text{orientations}) dr^{\alpha \beta}.
\]
The quantity in the brackets is easily identified as the ensemble angular average of the scalar product \((R_{12} \cdot \nu)\), so that finally we obtain
\[
\frac{\beta P}{\rho} = 1 - \frac{2\pi \beta \rho}{3} \sum_{\alpha, \beta} \int g^{\alpha \beta}(r^{\alpha \beta}) \frac{du^{\alpha \beta}}{dr_{12}^{\alpha \beta}} ((R_{12} \cdot \nu^{\alpha \beta})(r^{\alpha \beta})^2 dr^{\alpha \beta}.
\]
Eqn. (3.55) is completely general and holds for any site-site interaction models. In this form it is especially convenient for the computer simulations; for theoretical considerations it can be rewritten into other forms, e.g., using spherical harmonics (for other expressions and details see the original paper [38]).
Chapter 4

Pseudoexperimental techniques

Making use of a rare computing facility available that time, MANIAC computer in the Los Alamos National Laboratory in New Mexico, Metropolis and his colleagues[8] carried out in 1953 a first computer experiment on a system of interacting particles. They considered a model fluid system made up of 64 hard spheres (billiard balls in three dimensions) and determined its equation of state and the radial distribution function. Although this provided molecular physicists with an experimental tool enabling them for the first time to test and verify their methods and predictions, future extensions and applications were not on the agenda. As remarked 50 years later by one of the members of the Los Alamos team, M. Rosenbluth, in his talk given at a conference celebrating 50th anniversary of the Metropolis algorithm[39], there was a toy available so why not to use it for various 'scientific games', and this was just one of them. Nothing more, no clear idea and intentions for further development. And yet 15 years later, in the late 1960’s, the role of such computer experiments was not yet quite clear, their potential still not envisaged, and those who were experimenting with these computer simulations were not fully accepted by the molecular physics community; it was suggested that they be called machinests, because they were just operating machines, nothing more. Only the fast development of computers, their power and speed, and their general availability opened eyes of scientific community, and not only of theoreticians dealing with molecules and molecular systems, but mainly applied scientists and engineers have found a tool to model and study problems otherwise defied.

In general, the terms ‘computer experiment’ or ‘computer simulation’ may mean the use of computers to get a solution of a complex set of equations describing the problem at hand. Applications may thus be so disparate as studies of the explosion of stars, flow of liquids in complex tubes, growth of crystals, the structure of the phase interface, technological processes etc. In this textbook we will try to avoid the use of the above terms and will adhere to the term ‘molecular simulations’ to explicitly express the fact that we deal with systems of a huge number of molecules. The considered systems are thus subject to the laws of statistical physics and the properties we are interested in are, e.g., thermodynamic properties and structure of pure
fluids and their mixtures, phase equilibria, etc. Moreover, molecular simulations make it also possible to obtain properties otherwise inaccessible by classical experiments (e.g., the structure of the interface, molecular re-arrangement at the level of picoseconds, or processes at explosions), or to study molecular systems at experimental conditions hardly reachable on the Earth (e.g., extremely high pressures and/or temperatures existing in the interior of stars). For all these reasons they may be viewed as experiments; in fact, they generate data (referred to usually as pseudoexperimental data) for systems which do not exist in reality. On the other hand, some researchers view molecular simulations as exact theory because, after all, they provide a correct solution of the problem at hand.

We have already mentioned that the modern theory of liquids has come to existence thanks to pseudoexperiments (computer experiments). Without them we would hardly discover the molecular mechanisms governing the macroscopic behavior of macroscopic systems (see the next Chapter). Nowadays molecular simulations have become a routine tool to study a large variety of problems and it is therefore evident that thousands of results obtained by simulations have been reported in literature every year with this number ever increasing. For detailed and fundamental knowledge of simulations we refer the reader to some of existing monographs, e.g. [29, 30]. This chapter should thus serve only as a simple introduction to molecular simulations providing their basic principles.

In equilibrium statistical mechanics of interacting particles the primary quantity we have to compute is the configurational partition function from which all other properties can be obtained. In fact, what we always calculate is an average of a quantity \( X \). There are two types of averages. One follows the time development of the system and quantity \( X \) is then averaged over time,

\[
\overline{X} = \frac{1}{\tau} \int_0^\tau X(t) \, dt .
\]  

(4.1)

The other possibility to average \( X \) follows the walk of the system in the configuration space and the average is thus evaluated as

\[
\langle X \rangle = \frac{1}{k} \sum_{i=1}^k X_i ,
\]  

(4.2)

where \( X_i \) is the value of \( X \) at configuration \( i \). Provided that the system is ergodic then, according to the ergodic hypothesis, both average values are identical in the thermodynamic limit.

In numerical computations we have always to work with discretized variables which means that the time average (4.1) is actually computed as

\[
\overline{X} = \frac{\Delta t}{\tau} \sum X(t_i) .
\]  

(4.3)

Both ways thus do in fact the same, average values \( X_i \), and they differ only in the way how they generated configurations. In molecular simulations we generate these two types of averages.
4.1 Basics of molecular dynamics

Molecular dynamics follows the time development of the system that is governed (in the case of classical mechanics) by Newton’s equations of motion,

\[
\frac{d^2 \mathbf{R}_i}{dt^2} = \frac{1}{m_i} \mathbf{F}_i ; \quad i = 1, \ldots, N ,
\]

where \( \mathbf{F}_i \) is the force acting on molecule \( i \) of mass \( m_i \). For simplicity we will further assume that \( m_i = m \) for all \( i \). This set of equations must be supplemented by initial conditions, i.e., by the positions, \( \mathbf{r}_i(t = t_0) \), and velocities, \( \mathbf{v}_i(t = t_0) \equiv (d\mathbf{r}_i/dt) \) at some initial time \( t_0 \). Using numerical methods we cannot get the solution (trajectory) at any moment \( t \) but only at discrete values \( t, t + h, t + 2h, \ldots, t_{\text{max}} \) where the time step \( h \) is sufficiently small so that the changes in positions and of forces are ‘small’. The actual choice of \( h \) depends then on the system at hand, numerical method used, etc.

4.1.1 General remarks

Provided that the force is continuous, the trajectories are smooth functions that can be expanded in a Taylor series,

\[
\begin{align*}
\mathbf{r}^{(p)}(t + h) &= \mathbf{r}(t) + \frac{d\mathbf{r}}{dt}|_{t'=t}h + \frac{1}{2} \frac{d^2\mathbf{r}}{dt^2}|_{t'=t}h^2 + \frac{1}{6} \frac{d^3\mathbf{r}}{dt^3}|_{t'=t}h^3 + \ldots \\
&\equiv \mathbf{r}(t) + \mathbf{v}(t)h + \frac{1}{2} \mathbf{a}(t)h^2 + \frac{1}{6} \mathbf{b}(t)h^3 + \ldots 
\end{align*}
\]

\[
\begin{align*}
\mathbf{v}^{(p)}(t + h) &= \mathbf{v}(t) + \mathbf{a}(t)h + \frac{1}{2} \mathbf{b}(t)h^2 + \ldots \\
\mathbf{a}^{(p)}(t + h) &= \mathbf{a}(t) + \mathbf{b}(t)h + \ldots \\
\mathbf{b}^{(p)}(t + h) &= \mathbf{b}(t) + \ldots ,
\end{align*}
\]

where by superscript \( (p) \) we have indicated that the quantities at time \( t + h \) are predicted from those at time \( t \). Thus, formally, from the knowledge of \( \mathbf{r} \) at time \( t \) we may get \( \mathbf{r} \) at the next moment \( t + h \), from this solution then the solution at time \( t + 2h \), and so on. And the same applies to velocities and accelerations. However, this cannot be a good idea because it does not make use of the equation of motion at all. Consequently, the result of such a straightforward method starts very quickly to deviate from the correct solution as time advances.

A better but yet a very simple method to improve the above scheme is that of Euler that incorporates, as it should be, into the solution the equation of motion. Specifically, from the equation of motion we know that the acceleration is given by forces which are determined by the positions of the particles,

\[
\mathbf{a}(t) = \frac{1}{m} \mathbf{F}[\mathbf{r}(t)] .
\]
Thus, truncating the above Taylor expansion, for simplicity, after the $h^2$ term and using (4.9) we get for the solution at time $t + h$

$$r^{(p)}(t + h) = r(t) + v(t)h + \frac{1}{2m}F[r(t)]h^2$$  \hspace{1cm} (4.10)

$$v^{(p)}(t + h) = v(t) + \frac{1}{m}F[r(t)]h$$ \hspace{1cm} (4.11)

$$a^{(p)}(t + h) = \frac{1}{m}F[r(t)]$$ \hspace{1cm} (4.12)

Although the Euler method is not so bad, its accuracy is not sufficient. To devise a better (more accurate and fast) method to solve the set of Newton’s equations we may pursue the following idea. We may view the above Taylor expansion as an extrapolation from time $t$ to time $t + h$. Within this interpretation, we may expect that a better solution we get if more ‘history’ is used, i.e., we use the knowledge of solutions at more previous times than at only one previous moment. But we may also go a bit further. From the known forces at time $t$ we predict, as usual, accelerations $a^{(p)}$ at time $t + h$, see Eqns. (4.12). But simultaneously, from Eqn. (4.10) we also predict trajectories at time $t + h$, $r^{(p)}(t + h)$, that determine forces and hence, via the equations of motions, also accelerations at time $t + h$, $a^{(e.m.)}(t + h)$. Now we have two values for $a(t + h)$: one from the direct prediction, $a^{(p)}$, and the other, $a^{(e.m.)}$, obtained from the predicted trajectories. Both values should be the same, provided that the solution is correct, but this will not be the case. Their difference may thus serve as a measure of inaccuracy but, primarily, it may serve to correct the predicted solution. This is the principle of predictor-corrector methods.

Readers familiar with numerical analysis may ask the question why do we bother here about methods to solve ordinary differential equations, or even why do we talk about molecular dynamics as a special discipline, when libraries are filled with textbooks on numerical solutions of differential equations. The reason is very simple, it is the right-hand side of Eqn. (4.4). To be specific, the force acting on molecule $i$ in the absence of any external field is

$$F_i = -\frac{\partial U(1, \ldots, N)}{\partial r_i} = - \sum_{k=1, k \neq i}^{N} \frac{\partial u(r_{ik})}{\partial r_i} = - \sum_{k=1, k \neq i}^{N} \frac{du(r_{ik})}{dr_{ik}} r_{ik}.$$ \hspace{1cm} (4.13)

The evaluation of the right-hand side of one equation requires thus evaluation of $N$ terms, of the entire set of $N$ equations then $N^2$ terms. For $N \approx 10^3$ it means the evaluation of $10^6$ terms in only one single step. If we further realize that evaluation of pair energies themselves may be time consuming it becomes evident that all methods that require more than one iteration step (one evaluation of the right-hand side) to get from $t$ to $t + h$ are practically disqualified (provided that we want to get a solution before we retire), which applies particularly to Runge-Kutta methods.
4.1. BASICS OF MOLECULAR DYNAMICS

4.1.2 Verlet method

As mentioned above, one may improve the solution at time \( t + h \) (i.e., the extrapolation to time \( t + h \)) if more ‘history’ is used, it means, to make use of the positions and velocities also at previous times \( t - h, t - 2h, \ldots \). A simple implementation of this idea was proposed by Verlet [40]. He uses the knowledge of the trajectory at only two previous times, but in a quite smart way.

Expanding \( r(t \pm h) \) about \( r(t) \) to the second order we may write

\[
\begin{align*}
    r(t + h) &= r(t) + v(t)h + \frac{1}{2}a(t)h^2, \\
    r(t - h) &= r(t) - v(t)h + \frac{1}{2}a(t)h^2.
\end{align*}
\]

(4.14) (4.15)

By adding these two equations together we get:

\[
    r(t + h) = 2r(t) - r(t - h) + a(t)h^2.
\]

(4.16)

Finally, after making use of the equation of motion at time \( t \) we get a recurrent formula

\[
    r(t + h) = 2r(t) - r(t - h) + \frac{F[r(t)]}{m}h^2,
\]

(4.17)

enabling us to determine the location of particle \( i \) at time \( t + h \) from the knowledge of its location at times \( t - h \) and \( t \). This is the basic form of the Verlet algorithm with error of the order \( \mathcal{O}(h^4) \) (the first neglected term is proportional to \( h^4 \)).

To implement the method, in addition to the initial conditions at time \( t_0 \) we need also positions at time \( t_0 - h \). This however poses no problem; to get it for this purpose we may use again the Taylor expansion,

\[
    r_i(t_0 - h) = r_i(t_0) - hv_i(t_0) + \frac{h^2}{2m}F_i(t_0).
\]

(4.18)

As we see, in this formulation the method does not operate with velocities at all. Although they are not needed for the time evolution, their knowledge is required, e.g., to compute the kinetic energy whose evaluation is necessary to test the conservation of the total energy. One can compute the velocities directly from positions at times \( t + h \) and \( t - h \),

\[
    v(t) = \frac{r(t + h) - r(t - h)}{2h},
\]

(4.19)

whose error are however of order \( \mathcal{O}(h^2) \) rather than \( \mathcal{O}(h^4) \). To overcome this difficulty, other variants of the Verlet algorithm have been developed.

Despite its simplicity, Verlet’s method has turned out to be surprisingly very successful and has become the most common molecular dynamics method. The reason behind its success, besides the simplicity, is its time reversibility which leads to an excellent energy conservation during simulation.
CHAPTER 4. PSEUDOEXPERIMENTAL TECHNIQUES

4.1.3 Predictor-corrector algorithms

As time evolves, due to accumulation of errors the predicted trajectories deviate more and more from the correct trajectories. The idea of the predictor-corrector methods is to suppress the propagation of errors. In the case of Newton’s equations of motion we have at time $t + h$ only one measure characterizing inconsistency of the solution: difference $\Delta a(t + h) = a^{(c,m)}(t + h) - a^{(p)}(t + h)$. This difference is used to correct the predicted solution in one step only within the following general scheme:

$$
\begin{align*}
    r^{(c)}(t + h) &= r^{(p)}(t + h) + c_1 \Delta a h^2 \\
v^{(c)}(t + h) &= v^{(p)}(t + h) + c_2 \Delta a h \\
a^{(c)}(t + h) &= a^{(p)}(t + h) + c_3 \Delta a,
\end{align*}
$$

where superscript $(c)$ denotes the corrected solution, and $c_i$ are certain coefficients determined from the conditions of optimum stability and accuracy of the trajectories and are discussed in literature on differential equations[41].

The corrector-predictor method explained above is usually referred to as Gear integrator and represents an alternative to Verlet’s method.

4.1.4 Remarks and comments

To conclude this short account on the MD method a few remarks seem appropriate.

In general, it is necessary first to decide what method to choose from the two main available: the Verlet or Gear predictor-corrector method. This is rather a matter of the personal experience or taste than a principle physical problem. Second, it is evident that one would like to use the integration step, $h$, as large as possible, to obtain the long trajectory with the minimum computations. However, the large integration step will inevitably lead to large error. On the other hand, very small steps, in addition to an enormous increase of the computational effort, may lead to an increase of the effect of rounding errors that may accumulate. A related question is what criterion to apply.

In microcanonical ensemble (which is the case of the MD simulation considered so far) the total energy is conserved. The sum of the kinetic and potential energies (in the absence of an external field) must be therefore a constant. However, in numerical computations it will not be exactly a constant and its changes during the simulation are an important indicator. Thus, due to

(i) time irreversibility of the chosen integrator we will observe a systematic decrease or increase of the energy (the so called drift), and

(ii) statistical accumulation of errors (coming from various sources) we will observe noise.
4.2 BASICS OF MONTE CARLO METHOD

4.2.1 Naive versus smart sampling

Monte Carlo methods in science and engineering derive their name from the historical Mecca of gambling - Monte Carlo, where randomness determines gambler’s success or failure. The same type of randomness is used in mathematics, e.g., for evaluation of multidimensional integrals.

Let us consider a complex shaded plane figure F within a square S shown in Fig. 4.1a and try to evaluate its area $A$. The figure occupies a fraction of the square S and if we generate a sufficient number of points randomly scattered in the square and count those that fall into the shaded area then the fraction of the successful hits of this area with respect to the total number of attempts gives the fraction of the area occupied by the figure,

$$A_{\text{shaded area}} = \frac{\text{number of successful hits}}{\text{number of all attempts}} \times A_{\text{total area}}.$$ (4.23)
CHAPTER 4. PSEUDOEXPERIMENTAL TECHNIQUES

In equilibrium statistical mechanics of interacting particles the primary quantity we have to compute is the configurational partition function from which all other properties can be obtained. In the quasiclassical approximation the average value of any physical quantity \( X \) is then given by

\[
\langle X \rangle = \frac{1}{Q_N} \int X(\{q^N\}) e^{-\beta U(\{q^N\})} d\{q^N\},
\]  

(4.24)

where \( Q_N \equiv Q_{NVT} \) is the configurational integral, see Eqn. (2.10). It means, we have in fact to evaluate a multidimensional integral. Why do we then talk about Monte Carlo simulations as a special discipline when we could use a multidimensional variant of the method exemplified above? Simply because that it would not work. For obvious physical reasons.

To be specific, the situation we encounter in statistical mechanics can be exemplified again by evaluation of an area of a 2D figure, see Fig. 4.1b. Although the area \( A_F \) of figure F to be evaluated is very, very small, this is not the main problem. The problem is caused by the fact that the location of the figure (i.e., of configurations of the system around the equilibrium) within a much, much larger area \( A_S \) (the configuration space) is not known a priori. It means that, unlike the case of the previously considered evaluation of the 2D integral, we cannot circumscribe figure F within some well defined figure of a comparable area to make successful random hits quite probable. Inversely, even if we generated a huge number of points randomly scattered within square S, the probability that some of the points fall into figure F is (nearly) zero and we cannot thus say anything about the actual size of area \( A_F \).

Let us now translate this into the language of physics. In the integrand of Eqn. (4.24) there stands the Boltzmann factor representing the probability of occurrence of a configuration of particles \( \{q_1, \ldots, q_N\} \) with the total energy \( U(q_1, \ldots, q_N) \). If we now place particles into the simulation box randomly (the so called ‘naive Monte Carlo’), then nearly with certainty we find at least two particles very close one to another. Since at such small separation the repulsive force is very strong, we will get \( \exp[-\beta u] \approx 0 \); hence the probability of occurrence of the configuration will be zero. Using this naive route we may thus hardly generate sufficient number of configurations with high probability of their occurrence.

How to overcome this problem? Let us assume that we manage to construct a configuration that may potentially exist, i.e., \( \exp[-\beta U(\{q^N\})] \neq 0 \). Instead of generating a new configuration from the scratch we may use the existing configuration and only slightly change it (for instance, by displacing one of its particles) so that \( \exp[-\beta U(\{q^N\})] \) will be non-zero again. We may proceed in this way and generate configurations one after another; i.e., in the language of Fig. 4.1b, to move within vicinity of or even within figure F. The only problem is that these changes of configurations must be made within the spirit of the laws of physics.

To summarize, instead of trying to generate new configurations again and again from the scratch (naive MC), we always use the existing configuration to gradually move through the space of configurations to sample efficiently the regions that give significant contributions to the
4.2. BASICS OF MONTE CARLO METHOD

integral to be evaluated (smart MC). The remaining problem is, in addition to the correctness of the moves, to show that the final result we get really is a correct estimate of the ensemble average (4.24).

4.2.2 Metropolis algorithm

Physical systems let by themselves (i.e., not being subject to external forces) evolve towards the equilibrium characterized by extremes of their thermodynamic functions, e.g., the minimum of energy. Thus, if we have a configuration with energy $U_{\text{now}}$ and the configuration after the change has a lower energy, $U_{\text{next}} < U_{\text{now}}$, then the change goes in the right direction because the probability of occurrence of the new configuration is greater than the previous one, $\exp[-\beta U_{\text{next}}] > \exp[-\beta U_{\text{start}}]$, and the new configuration is accepted (can be added to the archive of generated configurations). But can we continue indefinitely along this scheme towards the states with lower energies? What happens when the system reaches an equilibrium? Does it come to a standstill? No! We know that the system continues to ‘live’, i.e., it fluctuates around the equilibrium which means that the energy may also increase and such configurations are thus legitimate and must also be accepted. What may be the rule of acceptance of such a fluctuation into a wrong direction? Using intuition (common sense) we may guess that the greater is the deviation, the less probable it should be to accept such a fluctuation.

Let us implement now this idea mathematically. We know that the ratio of the occurrence probabilities of the two states, the new one and existing one, is $e^{-\beta U_{\text{next}}} / e^{-\beta U_{\text{now}}}$. To satisfy this condition, the configuration with a higher energy must be accepted just with this probability because in the case of the change in the opposite direction (lower energy of the new configuration) the new configuration is always accepted. Thus,

$$P = e^{-\beta(U_{\text{next}} - U_{\text{now}})}.$$  \hspace{1cm} (4.25)

The criterion of the acceptance of the trial move may thus be written in a compact form as

$$P_{\text{acc}} = \min\{1, e^{-\beta(U_{\text{next}} - U_{\text{now}})}\},$$  \hspace{1cm} (4.26)

or

$$P_{\text{acc}} = \min\{1, \frac{\pi_{\text{now}}}{\pi_{\text{next}}}\}. \hspace{1cm} (4.27)$$

This quite simple method was devised by Metropolis and his colleagues[8] and has become the most common scheme of MC simulations (it means that other possibilities of generating configurations exist as well). There are now two questions that arise: (i) how to implement the above criterion, and (ii) whether the obtained result really is the correct ensemble average. The answer to the former question is only a simple technical problem. As for the latter, to prove the correctness of the method, the theory of Markov chains is used but this is out of the scope of this book and we refer the reader to textbooks on simulations for further details.
Chapter 5

Basic theoretical tools

There are three basic methods used to develop a full description of fluids at the molecular level: (i) expansions in powers of an appropriate parameter, typically density, (ii) integral equation theories aiming at finding the pair correlation function for the given interaction model $u(q_1, q_2)$, and (iii) perturbation theories building up the theory by means of developing corrections over the results for a simpler system. To develop these methods, in addition to common mathematical operations, two specific mathematical techniques are used: functional differentiation and diagrammatic technique. The former not only introduces formally the correlation functions but makes it also possible to perform certain operations on graphs which, in combination with the latter forms the basis of the topological reduction method.

The first of the above methods has as its target the thermodynamic properties and makes it possible to derive also various approximations for the integral equations. The correlation functions (of any order) can also be obtained from the integro-differential equations (Bogoljubov-Born-Green-Kirkwood hierarchy [20]) but this route is a bit cumbersome not providing an analytic solution, neither for the simplest models, and was used in the early stage of the development of theories of fluids. However, these equations play the central role for non-equilibrium fluids [11]. Since in the case of the perturbation theories the full description of the reference (simple) system must be known, including the pair correlation function, these theories could hardly be developed, as we will see later, independently of integral equation theories. Also, specific methods to describe the properties of simple (reference) systems can be used as, e.g., the scaled particle theory.

5.1 Functional differentiation

In statistical physics the physical properties of many body systems are typically given by integrals, e.g., the partition function and derived properties, which are thus functionals. A typical example are the correlation functions which are functionals of the intermolecular potential.
5.1. FUNCTIONAL DIFFERENTIATION

Functionals we encounter in this text may be written as

\[ F \equiv F[y(t)] = \int_{a}^{b} k(t)y(t) \, dt , \quad (5.1) \]

with \( k(t) \) and \( y(t) \) being differentiable functions on the interval \((a, b)\). From the mathematical point of view, the functional differentiation is a simple extension of the differentiation of functions. Whereas in the differentiation of a function \( f(x) \) we consider a change of \( f \) with respect to an infinitesimal change of the independent variable \( x \), \( \Delta f / \Delta x \), in the functional differentiation it is the change of the functional with respect to a change of function; thus, if \( F \) is a functional of \( y(t) \), it is \( \delta F / \delta y \), where we denote the differentiation of the functional by \( \delta \) to distinguish it from the common differentiation of functions. For details and full mathematical definitions see, e.g., [42, 19, 20].

Functional differentiation is used in theories of fluids (statistical physics) to formally introduce the correlation functions. The power of this elegant method lies in that it makes possible to provide mathematical operations on graphs (topological reduction, see the next section) and derive various relations between the correlation functions.

We will use the grandcanonical ensemble and consider a system in an external field. Thus, the total energy, \( U \), of the system with \( N \) particles is now

\[ U^{(\text{tot})}(\{q^N\}) = U(\{q^N\}) + \sum_{i=1}^{N} \Psi(q_i) , \quad (5.2) \]

where \( \Psi(q_i) \) is the energy of particle \( i \) due to the presence of the external field. Defining the activity \( a \),

\[ a(q_i) = ze^{-\beta\Psi(q_i)} , \quad (5.3) \]

the partition function can then be written in the form

\[ \Xi = \sum_{N \geq 0} \frac{1}{N!} \int e^{-\beta U(\{q^N\})} \prod_{i=1}^{N} a(q_i) \, dq_1 \ldots dq_N . \quad (5.4) \]

For the change of the partition function \( \Xi \) due to a variation of the activity we then get

\[ \delta \Xi = \sum_{i \geq 1} \frac{N}{N!} \int \left[ \int e^{-\beta U(\{q^N\})} \prod_{i=2}^{N} a(q_i) \right] \delta a(1) \, dq_1 , \quad (5.5) \]

where we have explicitly indicated that all molecules give the same variation of the activity. Hence, the functional derivative of \( \Xi \) with respect to \( \delta a \) assumes then the form

\[ \frac{\delta \Xi}{\delta a(q_1)} = \sum_{N \geq 1} \frac{1}{(N-1)!} \int e^{-\beta U(\{q^N\})} \prod_{i=2}^{N} a(q_i) \, dq_2 \ldots dq_N . \quad (5.6) \]
Comparing now this expression with Eqn. (3.8) shows that the first-order distribution function of the system in the external field is given by

$$\rho^{(1)}(q_1, \Phi) = a(q_1) \frac{\delta \ln \Xi}{\delta a(q_1)} . \tag{5.7}$$

The pair distribution function is obtained in a similar way,

$$\rho^{(2)}(q_1, q_2, \Phi) = a(q_1)a(q_2) \frac{1}{\Xi} \frac{\delta^2 \Xi}{\delta a(q_1)\delta a(q_2)} , \tag{5.8}$$

and, in general,

$$\rho^{(k)}(q_1, \ldots, q_k, \Phi) = \frac{1}{\Xi} \prod_{i=1}^{k} a(q_i) \frac{\delta^k \Xi}{\prod_{i=1}^{k} \delta a(q_i)} . \tag{5.9}$$

In addition to variations of the activity it is possible to consider also variations with respect to other functions defining the partition function. Considering the variation of \( \Xi \) with respect to the external potential we get:

$$\delta \ln \Xi = \frac{1}{\Xi} \sum_{N \geq 1} \frac{z^N}{(N-1)!} \int \left[ \int \prod_{i \geq 1, j > i} [1 + f(q_i, q_j)] \prod_{i \geq 1} a(i) \, dq_i \ldots dq_N \right] \delta[-\beta \Psi(q_1)] \, dq_1 . \tag{5.10}$$

Now, as above, from the definition of the distribution of the first order it follows that

$$\rho^{(1)}(q_1, \Phi) = \frac{\delta \ln \Xi}{\delta(-\beta \Psi(q_1))} . \tag{5.11}$$

Similarly, we may consider the variation due to the Boltzmann factor, \( f(q_1, q_2) = \exp[-\beta u(q_1, q_2)] - 1 \). Since

$$e^{-\beta U(q^N)} = \prod_{i \geq 1, j > i} [1 + f(q_i, q_j)] , \tag{5.12}$$

we then get

$$\delta \Xi = \sum_{i \geq 2} \frac{N(N-1)}{2N!} \int \left[ \int \prod_{i \geq 1, j > i} [1 + f(q_i, q_j)] \prod_{i \geq 1} a(i) \, dq_i \ldots dq_N \right] \delta f(q_1, q_2) \, dq_1 \, dq_2 , \tag{5.13}$$

and hence,

$$\rho^{(2)}(qv_1, q_2, \Psi) = 2 \frac{\delta \ln \Xi}{\delta f(q_1, q_2)} [1 + f(q_1, q_2)] . \tag{5.14}$$

Above we have shown an alternative way to introduce the correlation functions. As it is seen, it is an elegant and powerful tool but we always have to bear in mind that the above derivations are formal and purely mathematical and without any physical content; thus, without an a priori knowledge what we want to get we would hardly introduce the correlation functions in this way.

As an example we give here the definition of a function

$$c(q_1, q_2) = \beta \frac{\partial \phi_1}{\partial n^{(1)}(q_2)} + \frac{\delta(q_1, q_2)}{n^{(1)}(q_1)} , \tag{5.15}$$
5.2. DIAGRAMMATIC TECHNIQUE

which is known under the name the direct correlation function. Using now the relation

\[ \int \frac{\partial n^{(1)}(q_1)}{\partial \phi(q_3)} \frac{\partial \phi(q_3)}{\partial n^{(1)}(q_2)} dq_2 = \delta(q_1, q_2), \]

we get an equation

\[ h(q_1, q_2) = c(q_1, q_2) + \rho \int c(q_1, q_3) h(q_2, q_3) dq_3, \]

introduced and discussed later in this Chapter. Other examples, definitions, and relationships along with further details on the functional differentiation can be found in, e.g., [20] or [19] to which readers are referred.

5.2 Diagrammatic technique

To represent integrals defining the configuration integral, Eqn. (2.30), by means of graphs goes back to and details of his method can be found in [1]. Further, and quite a different development is associated with the functional differentiation and topological reduction introduced by Morita [43] and used, e.g., to derive various closures for the Ornstein-Zernike integral equation (see the following Section) or in the thermodynamic perturbation theory.

Let us consider the \( N \)-particle configurational integral \( Q_N \), Eqn. (2.30). There are \( N(N - 1)/2 \) factors in the product (2.30) and thus altogether \( 2^{N(N-1)/2} \) terms, calling diagrams, contributing to the sum. It is evident that to handle with this enormous number of terms is a formidable task, not to mention their evaluation. The diagrammatic technique pictures the individual integrals as graphs and makes it possible to manipulate with them. To understand this technique and its essence, let us begin first with two simple examples related to the original work of Mayer.

A graph associated with a diagram is defined as follows:

*Each integration variable (particle) is pictured as a filled circle (field point) and the function \( f_{ij} \) as a line connecting circles \( i \) and \( j \).*

Thus, for instance:

\[ \int \ldots \int d\{7\} f_{12} f_{14} f_{23} f_{56} = \]

A line connecting circles \( i \) and \( j \) is called a bond. Other two important terms we will need are (i) path and (ii) connectivity. Both terms are self-explanatory. If there is a path between
any pair of circles \(i\) and \(j\) (i.e. a set of bonds \(i - k, k - l, \ldots, m - j\)) of the given diagram, then the diagram is called connected or simply a cluster.

In general, field points may carry also a function; in the above example it is only the integration variable (not indicated). Since the integration cannot depend on the labeling of the integration variables, many diagrams will contribute the same value and they will differ only in the assignment of integration variables to the field points. It is thus useful to assign to the field points and arbitrary labeling and add to the diagram the combinatorial factor. This is the typical procedure introduced by Mayer and also typical for derivations in the canonical ensemble.

To demonstrate the manipulation with diagrams we will give below two examples.

We see that the diagram in (5.18) consists of three clusters (single circle is considered as a connected diagram), and this is a general property of all diagrams appearing in (2.30): Any diagram is made up of independent clusters. Since there are no connections between the clusters in the diagram, the resulting integral factors into the product of integrals, each of them corresponding to one cluster. Thus, considering again the diagram in (5.18):

\[
(4)(1) (2) (3) (6) = \int \ldots \int d(1) \ldots d(4)f_{12}f_{14}f_{23} \int d(5)d(6)f_{56} \int d(7) \quad (5.19)
\]

The above example may seem trivial, but the usefulness of picturing the multidimensional integrals by means of graphs will become clear soon. To evaluate the configurational integral we need, in general, to evaluate all possible clusters. A general idea to handle with diagrams making up the configurational integral is (i) to start with a cluster of some specific property (called the core cluster), (ii) collect into one group all such clusters and evaluate the common factor given by the contribution of all the remaining clusters in the given diagram, and then (iii) to sum up over all groups of core clusters. This idea may be exemplified by considering the following simple example.

**EXAMPLE 1.**

Let us consider an ensemble of 4 particles. From all 64 terms making up the configurational integral \(Q_4\) let us collect together all diagrams containing 3-particle clusters. These diagrams are shown in Fig. 5.2, where we labeled the particle at the left lower corner as particle 1; labeling of other particles is immaterial but it is kept fixed. The diagrams are arranged in such a way that in one column there are all diagrams containing the 1-particle cluster of the same particle. If we sum up diagrams in columns it is evident that for each column we get the same result (if we assigned numbers to the remaining three vertices, then the diagrams in different columns would differ only in labeling of vertices). Further, we see that all clusters in one column are multiplied
5.2. DIAGRAMMATIC TECHNIQUE

Figure 5.1: 4-particle diagrams containing 3-particle clusters

by the same factor [see item (ii) above] ; in the considered case it is \( \int d(j) \) which is in fact the 1-particle configurational integral. It may be therefore convenient to evaluate the clusters in one column together:

The contribution of the clusters of the same order to \( Q_N \) is called the group integral (i.e., the above four clusters define the ternary group integral, \( b_3 \)). In general, we define the group integrals \( b_{l+1} \) as follows:

\[
b_{l+1} = \frac{1}{(l+1)!} \sum^{(l+1)} \int \cdots \int d\{l+1\} \prod' f_{ij},
\]

where \( \sum^{(l+1)} \) means sum over all clusters of \( l \) particles plus particle 1, the prime means that not all possible pairs \( i - j \) appear in the product, and the combinatoric factor has been introduced for convenience. It is important to keep in mind that we explicitly labeled one particle which means that the fourth column must be considered separately; it will contribute to the 1-particle group integral! Making now use of definition (5.20),...
we may write the total contribution from the first three columns (i.e. from the clusters of particle 1 plus two other arbitrary particles) as

$$3Q_1 \sum_l (2+1) \int \ldots \int \prod f_{ij} \frac{(N-1)!}{l!(N-l-1)!} Q_{N-l-1}(l+1)! b_{l+1}, \quad N = 4, \ l = 2 \quad (5.21)$$

where the fraction stands for the number of ways of choosing $l$ particles out of the $(N-1)$ particles.

The above simple example suggests that the systematic re-arrangement of diagrams in (2.30) may proceed as follows. We choose an arbitrary particle and label it as particle 1. Further, we pick up $l$ other arbitrary particles and consider all clusters formed by these $(l+1)$ particles. Since each of these clusters will bear the same factor, namely the sum of all possible diagrams formed by the remaining $(N-l-1)$ particles which is the configurational integral $Q_{N-l-1}$, these $(l+1)$-particle clusters will contribute to $Q_N$ a term

$$\frac{(N-1)!}{(N-l-1)!} (l+1)! b_{l+1} Q_{N-l-1}. \quad (5.22)$$

To get the entire $Q_N$ it remains only to sum up over all $l$:

$$Q_N = \sum_{l=0}^{\infty} \frac{(N-1)!}{(N-l-1)!} (l+1)! b_{l+1} Q_{N-l-1}, \quad (5.23)$$

and hence

$$\frac{Q_N}{Q_{N-1}} = \sum_{l=0}^{\infty} \frac{(N-1)!}{(N-l-1)!} (l+1)! b_{l+1} \frac{Q_{N-l-1}}{Q_{N-1}}. \quad (5.24)$$

This is a typical result for the process of forming groups of particular diagrams and summing then over all groups or of appending some diagrams to a core cluster.

We will now demonstrate another possible reorganization of the diagrams constituting $Q_N$. Let us consider a cluster in Fig. 5.2.
5.2. DIAGRAMMATIC TECHNIQUE

Keeping particle (4) fixed, we may place the origin of the coordinate system at this particle and integrate over \( d(5) \) independently of particles (1), (2), and (3). In other words, the cluster may be split into the product of two subclusters,

\[
\int \ldots \int d\{5\} f_{12} f_{14} f_{23} f_{34} f_{45} = \frac{1}{V} \int \int \int \int d\{1\} \ldots d\{4\} f_{12} f_{14} f_{23} f_{34} \int d(4) d(4/5) f_{45}, \quad (5.25)
\]

where \( d(4/5) \) means the integration over the relative coordinates of particle 5 with respect to particle 4 (recall also that the extra integration over \( d(4) \) in the second integral will produce the factor \( V \)). Particle (4) has thus the property that *snipping the cluster at this point separates the cluster into two subclusters*. Such a point we call a *node*. Whenever a node occurs in the cluster, the subclusters joined at this point factor into independent integrals. Because of this property such cluster is called *reducible* - it can be reduced to simpler units; clusters without any node are called *irreducible*.

The introduced properties of diagrams enable us to reorganize now the diagrams in (2.30): first, instead of choosing for the core cluster any connected diagram (cluster), we choose now the clusters with no more than one nodal point (the so called *star* clusters); then, in the subsequent step, we remove from the core the clusters with the nodal point ending up with expressions containing only irreducible clusters. Details of these manipulations can be found in nearly every textbook on statistical physics and we will not repeat them here and present only the final result: the virial expansion of the compressibility factor in powers of density with the diagrammatic expression of the virial coefficients:

\[
\frac{P}{\rho k_B T} = 1 - \sum_{i \geq 1} \frac{i}{i + 1} \beta_i \rho^i. \quad (5.26)
\]

Or, written in a more common way,

\[
\frac{P}{\rho k_B T} = 1 + \sum_{i \geq 1} B_{i+1}(T) \rho^i, \quad (5.27)
\]

where

\[
B_{i+1} = -\frac{i}{i+1} \beta_i, \quad i = 1, \ldots . \quad (5.28)
\]

are virial coefficients. The third and fourth virial coefficients in terms of the irreducible diagrams \( \beta_i \) are given as follows:

\[
B_3 = \frac{1}{3V} \quad (5.29)
\]
The other mentioned technique, the topological reduction, was introduced by Morita [43] to handle another type of diagrams. The technique is based on several lemmas enabling one also to differentiate the diagrams which leads not only to the expressions of sums of certain diagrams by means of a smaller number of more compact diagrams as above. Thus, all functions of interest are again expressed as an infinite series in powers of a parameter but, in addition, it makes it also possible to derive relationships between these functions. All these techniques lie outside the scope of this text and below we will present only the main results and otherwise refer the reader to specialized texts, e.g., [44].

Unlike the configuration integral which is a number, the multidimensional integrals defining the correlation functions are functions and must be therefore represented by another type of diagrams. These diagrams will contain two points, called root points and are depicted by open circles, representing variables $q_1$ and $q_2$. For example, the 5-particle cluster with two root points may look as follows:

\[
\begin{array}{c}
\text{(1)}
\end{array}
\quad \begin{array}{c}
\text{(2)}
\end{array}
\]

where we use, for simplicity, the notation $(i) \equiv q_i$.

The starting point for the further discussion is the expression of the background correlation function, $y$, by means of diagrams:

\[
y(R) = 1 + \begin{array}{c}
\text{(1)}
\end{array} + \begin{array}{c}
\text{(2)}
\end{array} + \begin{array}{c}
\text{(1)}
\end{array} + \begin{array}{c}
\text{(2)}
\end{array} + \begin{array}{c}
\text{(1)}
\end{array} + \begin{array}{c}
\text{(2)}
\end{array} + \begin{array}{c}
\text{(1)}
\end{array} + \begin{array}{c}
\text{(2)}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{c}
\text{(1)}
\end{array}
\end{array}
\quad \begin{array}{c}
\begin{array}{c}
\text{(2)}
\end{array}
\end{array}
\quad \begin{array}{c}
\begin{array}{c}
\text{(1)}
\end{array}
\end{array}
\quad \begin{array}{c}
\begin{array}{c}
\text{(2)}
\end{array}
\end{array}
\quad \begin{array}{c}
\begin{array}{c}
\text{(1)}
\end{array}
\end{array}
\quad \begin{array}{c}
\begin{array}{c}
\text{(2)}
\end{array}
\end{array}
\quad \begin{array}{c}
\begin{array}{c}
\text{(1)}
\end{array}
\end{array}
\quad \begin{array}{c}
\begin{array}{c}
\text{(2)}
\end{array}
\end{array}
\]

(5.30)
The existence of two kinds of points extends possibilities to distinguish (define) a larger amount of diagrams of specific properties. The simplest way to understand the properties of different diagrams is to view them as a current carrying network with the battery attached to the root points. Thus, for instance, the following diagram is called the *chain* (also nodal or series) diagram,

![Chain Diagram](image)

for the simple reason: it contains the nodal point(s) whose removal will cause that no current will flow between the root points. To be more accurate, in the case of diagrams with root points we define the node as such a field point whose removal leaves a disconnected diagram with root points lying in different components. An extension of the notion ‘nodal point’ is the introduction of an *articulation point*: it is a point whose removal (along with all associated bonds) leaves a disconnected diagram at least one component of which does not contain any root point. The graph without any articulation point is called *1-irreducible*. Similarly, an articulation pair of points in a connected diagrams is a pair whose removal leaves a disconnected diagram at least one component of which does not contain any root point. Such a diagram is called *2-irreducible*.

To continue in the classification, the following 4-particle diagram,

![Parallel Diagram](image)

is called the *parallel* diagram because the removal of any field point will not cause disconnection of current between the root points. A more precise definition says that the diagram is called parallel if there are at least two independent paths between the root points.

With these definitions we may now analyze the diagrams in 5.2) constituting the background correlation function. We see that the diagrams can be grouped into three types:

*chain diagrams* (the first three diagrams)

*parallel diagrams* (the fourth diagram)
elementary diagram (the last diagram)

The last diagram, represents the category of the most compact diagrams called, ironically, elementary. In general, the elementary diagrams are obtained by the parallel connection of two chain diagrams and connection of their field points by the $f$-bonds:

\[
(1) \circ (2) = (1) \circ (2)
\]

Denoting the sum of all chain diagrams as $N$, parallel diagrams as $P$, and that of the elementary diagrams as $B$, we may write

\[
y(1, 2) = 1 + N(1, 2) + P(1, 2) + B(1, 2),
\]

where $B$ is called the bridge function.

Now, to proceed further and give the reader a flavor of the method, we mention here one of the useful lemmas:

\[
L e t \Gamma \text{ be an infinite set of topologically different diagrams and } \mathcal{N} \text{ its subset containing only connected diagrams. If (i) any diagram from } \gamma \text{ can be constructed by diagrams from } \mathcal{N} \text{ and if (ii) for any diagram } D \text{ from } \mathcal{N} \text{ there exists a diagram in } \Gamma \text{ such that } D \text{ is one of its components, then } \\
\{ \text{sum of all graphs from } \Gamma \} = \exp[\text{sum of all graphs from } \mathcal{N}] - 1
\]

Using this and similar lemmas one can derive likely the most important result of the topological reduction, namely, the expression of the potential of the mean force by means of diagrams:

\[
W(1, 2) = B(1, 2) + N(1, 2). \quad (5.32)
\]

From the diagrammatic expression of the background correlation function it is straightforward to write a similar expression for the total correlation function where we now explicitly show the factor $\rho$ coming from the field points:
5.3. INTEGRAL EQUATIONS

\[ h(q_1, q_2) = + \rho \left\{ \begin{array}{c} \bullet \\
(1) \\
(2) \end{array} + \begin{array}{c} \bullet \\
(1) \\
(2) \end{array} + \begin{array}{c} \bullet \\
(1) \\
(2) \end{array} \right\} + \right\} + \right\} + \right\} + \right\} \]

\[ + \rho^2 \left\{ \begin{array}{c} \bullet \\
(1) \\
(2) \end{array} + \begin{array}{c} \bullet \\
(1) \\
(2) \end{array} + \begin{array}{c} \bullet \\
(1) \\
(2) \end{array} + \begin{array}{c} \bullet \\
(1) \\
(2) \end{array} + \begin{array}{c} \bullet \\
(1) \\
(2) \end{array} \right\} + \right\} + \right\} + \right\} + \right\} \]

Sorting out from these diagrams the nodal ones, we define a function \( c \):

\[ h(1, 2) = c(1, 2) + N(1, 2) . \]  

Further manipulations then show that \( h \) and \( c \) are related by another expression,

\[ h(q_1, q_2) = c(q_1, q_2) + c \rho \int c(q_1, q_3) h(q_2, q_3) dq_3 . \]  

It can be proved that function \( c \) is really identical with the function \( c \) introduced before and this equation is thus nothing else than the Ornstein-Zernike equation.

5.3 Integral equations

Integral equations for the pair correlation function may be introduced in a number of ways. Nonetheless, regardless which way we choose, in all cases we have to deal with the Ornstein-
Zernike (OZ) equation,

\[ h(r_{12}) = c(r_{12}) + \rho \int c(r_1)h(|r_1 - r_2|) \, d(r_2 - r_1) , \]  

(5.35)

where \( c(r) \) is the direct correlation function. An alternative form, used in numerical solutions and in discussions, reads as

\[ \gamma(r_{12}) = \rho \int c(r_1)[c(|r_1 - r_2|) + \gamma(|r_1 - r_2|)] \, d(r_2 - r_1) , \]  

(5.36)

where \( \gamma = h - c \).

The OZ equation can be "derived" by considering fluctuations in density, see, e.g., [20] or by the diagrammatic manipulations, see Eqn. (5.17). This equation was originally introduced by Ornstein and Zernike in 1914 [45] as the definition of another function, the direct correlation function, \( c \). The idea behind is to split the actual interaction acting on particle 1 in the dense system into the direct particle-particle interactions between particles 1 and 2 from a propagated effect, i.e., the effect of particle 2 on 1 through its effect on particle 3 which then affect particle.

Equation (5.35) contains two unknown functions, \( c \) and \( h \), and cannot be thus solved unless it is accompanied by another relation, called closure, between these two functions. One such a relation/approximation can be introduced practically immediately if we recall the intentional physical meaning of the direct correlation function, namely, to reflect the direct effect of the pair interaction in the system of interacting particles. As we know, at close interparticle distances the repulsive interaction sharply increases and the range becomes inaccessible. Many simple interaction models model thus the interaction at short interparticle separations by a hard wall. The pair correlation function in this region is thus zero,

\[ g(r) = 0 \quad \text{for } r \text{ within the core .} \]  

(5.37)

Outside the core we may use the physical interpretation of the function \( c \) and write

\[ c(r) = -\beta u(r) \quad \text{for } r \text{ outside the core .} \]  

(5.38)

These two equations, the first one of which is exact, define the so call mean spherical approximation (MSA) for models with a hard core. The MSA theory can be applied only to models with a hard core but this limitation is compensated but one important advantage: the MSA equation was solved analytically for several models and these results will be given and discussed in detail in subsequent sections.

In the preceding section we derived expressions for functions \( c, h, \) and \( W \) in the form of an infinite series in powers of density. Although we know the rules how to evaluate the coefficients of the series, only a few first term can be actually determined by means of the modern computers. However, the knowledge of these rule has another and very significant value: it enables us to find relations (and possibly some simplifications) among these functions.
5.3. INTEGRAL EQUATIONS

Altogether we have four equations,

\[
\begin{align*}
    h(1,2) & = N(1,2) + c(1,2) \quad (5.39) \\
    W(1,2) & = N(1,2) + B(1,2) \quad (5.40) \\
    h(1,2) & = e^{-\beta u(1,2)} + W(1,2) - 1 \quad (5.41) \\
    B(1,2) & = 0 \quad (5.43)
\end{align*}
\]

and the OZ relation, Eqn. (5.35), for five unknown variables. Whereas the correlation functions and the PMF represent physically important quantities, \( N(1,2) \) and \( E(1,2) \) are given by the sum of certain diagrams. The first elementary diagram contributing to \( W \) is of the order \( \rho^2 \) and it can be thus neglected in the low density region. If the pair interaction decreases rapidly with the intermolecular distance, the Mayer function becomes effectively zero beyond a certain distance. If the elementary diagrams are to give a significant contribution to \( W \), then the distance between all field points must be less than the limiting value. In other words, only a small part of the configurational space may contribute to the value of \( W \). Consequently, the diagram containing the \( f_{ij} \)-bond contributes a smaller value to the sum than the diagram which contains this bond. It seems therefore justified to set, as a plausible approximation,

\[
B(1,2) = 0 ,
\]

which implies that the direct correlation function is given by

\[
c(1,2) = h(1,2) - \ln[h(1,2) + 1] - \beta u(1,2) .
\]

Approximation (5.43) is called hyper-netted-chain approximation. Now we have four equations for four variables so that the set can be solved. The equation is too complex to allow to be solved analytically but numerical results are available for a number of potential models. In general, it should give better results for systems with short range interactions, for instance at high temperatures where the repulsive interactions dominate. On the other hand, its performance should deteriorate with decreasing temperature.

To derive the HNC approximation a certain classification of diagrams was employed. It is evident that this classification is not unique and another regrouping of the diagrams may be used. We will derive here now the second most important approximation, the Percus-Yevick one. This approximation was originally derived using different arguments and the following derivation serves thus simultaneously as an example of application of the diagrammatic technique without a direct physical significance.

Let us define functions \( M \) and \( D \) as follows:

\[
\begin{align*}
    c(1,2) & = f(1,2)[1 + N(1,2)] + M(1,2) \quad (5.45) \\
    D(1,2) & = M(1,2)e^{\beta u(1,2)} . \quad (5.46)
\end{align*}
\]
Now we may eliminate function $M$ which yields

$$c(1, 2) = f(1, 2)e^W(1, 2) + D(1, 2).$$

(5.47)

Setting now

$$D(1, 2) = 0,$$

(5.48)

we get an approximate expression for the direct correlation,

$$c(1, 2) = f(1, 2)e^W(1, 2).$$

(5.49)

It is easy to see that the condition $D = 0$ is equivalent to the condition $N = 0$ so that an alternative expression for the $c$ in the PY approximation is

$$c(1, 2) = f(1, 2)[1 + N(1, 2)].$$

(5.50)

The HNC and PY equations are two different approximations for the same function, the correlation function $g$. They differ in the approximation for the direct correlation function. It is easy to find that the difference is

$$c_{\text{HNC}} - c_{\text{PY}} = g \exp[\beta u] - \ln g - \beta u - 1.$$

(5.51)

For $r \to +\infty$, i.e., $u \to 0$, the above difference tends to zero. However, whereas $c_{\text{PY}}$ is exactly zero in the whole range of $u(r) = 0$, this need not be necessarily true for $c_{\text{HNC}}$; in the HNC approximation the direct correlation function may extend beyond the range of the interaction.

The HNC and PY integral equations have been solved, both analytically for some simple potential models, and numerically for a large set of interaction models. Detail and selected results will be discussed in detail in Section 8. General conclusions on their performance may be summarized as follows:

1. both the HNC and PY equations yield, over a large range of densities, much better results for the thermodynamic functions than equations based on the superposition approximation for $n^{(3)}$;

2. the PY equation performs better than the HNC equation in the case of steep repulsive potentials, especially for the model of hard spheres;

3. at low temperatures and moderate densities and for systems with attractive forces the HNC equation yields better results than the PY equation;

4. in the region of high densities the accuracy of both the HNC and PY equations considerably deteriorates.
5.4. PERTURBATION THEORIES

The HNC and PY approximations are the only approximations derived 'rigorously' using physical arguments. There have been numerous attempts to improve the performance of the integral equation, some of which being a modification of these approximations, other purely phenomenological. Probably the most spread of the latter attempts are the Martynov-Sarkisov and Verlet closures for the bridge function [46]:

\[ B_{MS}(r) = 1 + \gamma(r) - [1 + 2\gamma(r)]^{1/2} \]
\[ B_{V}(r) = \frac{\gamma^2}{2(1 + a\gamma)} , \]

where in the last equation the value \(4/5\) for parameter \(A\) was taken \textit{ad hoc}. Various modifications of both the Verlet and Martynov-Sarkisov closures have been proposed with a slight improvement over the original versions, see, e.g., [46]. The most successful modification of the integral equation approach is the 'Reference HNC' (RHNC) theory in which the bridge function of the investigated system is replaced by the bridge function of an appropriately chosen fluid of hard spheres [47].

5.4 Perturbation theories

A method of perturbation is likely the most common method used practically in all branches of science, including social sciences. The idea is to estimate the properties of the system/function of interest by means of those of another system/function whose properties are known. Accuracy of the final result may be then achieved by evaluating and adding some corrections. From the mathematical point of view it is an expansion in powers of a small parameter. For instance, in mathematics it is the Taylor expansion in powers of a small parameter in physics it may be the expansion of pressure in powers of density (virial expansion) or of the Helmholtz free energy in powers of \(\beta\delta u\) in the social science it may be the extrapolation of the behavior of the system to near future on the basis of its known contemporary behavior.

In molecular physics the primary independent quantity determining completely the system is the Hamiltonian. Thus, we would like to estimate the properties of the system with Hamiltonian \(\mathcal{H}\) by means of those of a system with Hamiltonian \(\mathcal{H}_0\). Adopting the assumption of pairwise additivity of interactions it means to determine the properties of the system with the intermolecular potential \(u(r)\) by means of the properties of the system with the potential \(u_0(r)\),

\[ u(r) = u_0(r) + \Delta u(r) , \]

where \(u_0\) is the reference potential and \(\Delta U\) perturbation. Formally, there is no problem to formulate such an expansion as it was done by Zwanzig [48] in the 1950’s. The only question is
CHAPTER 5. BASIC THEORETICAL TOOLS

what to choose for the reference potential $u_0(r)$ and what quantity may play the role of a ‘small’ parameter. The choice of the reference system is dictated by both mathematical and physical considerations. The main constraint imposed on the reference system is that its properties must be known to a high degree of accuracy and, preferably, in an analytic form. As one could guess, such a condition may be satisfied by very simple systems only whose Hamiltonian is far away from the original one. On the other hand, to make the perturbation corrections small to guarantee the fast convergence, the reference system should be as close as possible to the original one. However, in this case to describe the properties of the reference may be as difficult as dealing with the original system directly. For fluids no such a potential reference, even quite artificial, was hardly conceivable and as for a small parameter, density was used in the virial expansion about an ideal gas reference but this expansion is useless for dense fluids.

The above contradicting facts were thus likely the main and only reasons why the perturbation expansion was not used till the first pseudoexperimental results emerged. Dropping the mathematical arguments on the necessary smallness of the expansion parameter and using physical considerations it turns out that in fact it is not necessary to employ only a small parameter to get a good convergence which may be guaranteed by other means. To illustrate this statement, let us split the interaction potential, first rather arbitrarily, into two parts,

$$u(R) = u_0(R) + [u(R) - u_0(R)] = u_0(R) + \Delta u(R) ,$$

(5.55)

and use it in the expression for the internal energy, Eqn. (3.12):

$$U = 2\pi \rho \int g(R)u_0(R)R^2 \, dR + 2\pi \rho \int g(R)\Delta u(R)R^2 \, dR .$$

(5.56)

This expression is exact regardless of the value of the difference $\Delta u$. If we now find a system ‘0’ such that its structure is identical (very similar) to the structure of the original system,

$$g_0(R) \cong g(R) ,$$

(5.57)

then we may estimate the internal energy from

$$U \cong 2\pi \rho \int g_0(R)u_0(R)R^2 \, dR + 2\pi \rho \int g_0(R)\Delta u(R)R^2 \, dR = U_0 + \Delta U ,$$

(5.58)

where the last term plays the role of a correction to the reference system energy. In other words, the property of the liquid (dense system) can be determined by means of the perturbation expansion if the structure of the chosen reference system and the liquid at hand is the same.

It is now understandable why the perturbation theories began to be used only after molecular simulations methods were developed (along with availability of reasonably fast computers) during the 1960’s: It became possible to determine $u_0$ of any artificial system and to find a system with a similar structure and simultaneously simple enough to be able to determine its properties analytically. The perturbation approach to liquids was pioneered by Barker and Henderson,
5.4. PERTURBATION THEORIES

see their comprehensive review [49], followed by and many others. Here we present the basic possibilities for the expansion and equations; a detailed discussion on the choice of the reference system, evaluation of the correction terms and other technical details can be found in Chapter 9.

The perturbation expansion for fluids starts with the expansion of the Helmholtz free energy. Let us consider two systems defined by their total internal energies $U$ and $U_0$, resp., with $U = U_0 + \Delta U$. Thus,

$$-\beta F = \ln \left[ \int e^{-\beta(U+U_0)} \, d\{q^N\} \right].$$  \hspace{1cm} (5.59)

Expanding now the exponential function in powers of $\beta \Delta U$ we get

$$\beta(F - F_0) = \beta \langle \Delta U \rangle_0 - \frac{\beta^2}{2} \left( \langle \Delta U^2 \rangle_0 - \langle \Delta U \rangle_0^2 \right) + \ldots,$$  \hspace{1cm} (5.60)

where the angular brackets denote averaging over the configurations of the (reference) system ‘0’. The internal energy of the studied system is thus given by means of that of the system ‘0’ plus correction terms. For simplicity of the notation we will now consider henceforth only the simple fluids, i.e., the interaction between the molecules depends only on their separation. Assuming now the pairwise additivity of the intermolecular interactions then for the first correction term we get

$$\langle \Delta U \rangle_0/N = 2\pi \rho \int_0^{+\infty} g_0(r) \Delta u r^2 \, dr,$$  \hspace{1cm} (5.61)

where $g$ is the pair correlation function of the reference system ‘0’. To evaluate the first order correction for the internal energy we only need the pair correlation function. The second order correction is more complex and in addition to $g$ it requires also the knowledge of the correlation functions of the third and fourth order. In the 1960’s and 1970’s a lot of work was done on an approximate evaluation of this second order term for simple fluids but with the following development and diversion from theories towards simulations these results do not play any important role anymore. For interesting readers we refer them to monographs on theories of liquids, e.g., [20, 19].

To conclude this section we present an alternative way of deriving the correction terms. This method makes use of a coupling parameter, $\xi$, which makes it possible to change the effect of the perturbation:

$$U = U_0 + \xi \Delta U = U_0 + \xi \sum_{i<j} u(r_{ij}).$$  \hspace{1cm} (5.62)

From the change of the Helmholtz free energy with $\xi$,

$$\frac{\partial F}{\partial \xi} = \frac{1}{Q_N(\xi)} \int \Delta U e^{-\beta U_N(\xi)} \, d\{q^N\} = \langle \Delta U \rangle_\xi,$$  \hspace{1cm} (5.63)

we get

$$\beta F = \beta F_0 + \beta \int_0^1 \langle \Delta U \rangle_\xi \, d\xi.$$  \hspace{1cm} (5.64)
Assuming now again the pairwise additivity we get for the correction term the result

$$\langle \Delta U \rangle / N = -2\pi \rho \int g_0(R, \xi) \Delta u(R) R^2 \, dR ,$$

(5.65)

where now the pair correlation function depends on parameter $\xi$. This method, although in the end formally identical to the expansion presented above, is more general (less restrictive) and its application we will discuss in the Chapter 9.
Appendix A

Appendix: Basic thermodynamic relations

The thermodynamic state is defined by a set of independent variables and for every choice we have the corresponding thermodynamic potential. For simple systems with constant mass (number of particles, \(N\)), the independent variables are temperature, \(T\), volume, \(V\), pressure, \(P\), and entropy, \(S\), and the corresponding thermodynamic potentials are:

- Helmholtz free energy, \(A = A(V,T)\)
- Gibbs function, \(G = G(P,T)\)
- Enthalpy, \(H = H(S,P)\)
- Internal energy, \(U = U(S,V)\)

Starting from the second law of thermodynamics,

\[
dU = TdS - PdV ,
\]

the relations between these potentials are immediately derived:

\[
\begin{align*}
\text{d}A & \equiv \text{d}(U - ST) = -SdT - PdV , \\
\text{d}G & \equiv \text{d}(A + PV) = -SdT + VdP , \\
\text{d}H & \equiv \text{d}(G + ST) = TdS + VdP .
\end{align*}
\]

In these equations the changes in the potentials are total differentials which means that the coefficients at the differentials of the independent variables are their partial derivatives. For instance,

\[
\begin{align*}
S & = - \left( \frac{\partial A}{\partial T} \right)_V , \\
P & = - \left( \frac{\partial A}{\partial V} \right)_T , \\
V & = \left( \frac{\partial G}{\partial P} \right)_T .
\end{align*}
\]
Combining such derivatives with eqns. (A.1) - (A.4) one may get a number of further relations between various thermodynamic potentials. For instance, for the Helmholtz free energy and internal energy we get the relation

\[ U = A - T \left( \frac{dA}{dT} \right)_V = \frac{d(A/T)}{d(1/T)} = \frac{d(\beta A)}{d\beta}, \quad (A.8) \]

where \( \beta = 1/k_B T \) and \( k_B \) is the Boltzmann constant. Another useful relation is the expression for the dimensionless equation of state, the compressibility factor \( z \),

\[ z \equiv \frac{\beta P}{\rho} = \rho \frac{d(\beta A/N)}{d\rho}, \quad (A.9) \]

where \( \rho = N/V \) is the number density.

When the thermodynamic system is not closed, it means it can exchange the mass (particles) with its surrounding, a variable related to this change must be introduced. Thus, for changes at specified \( V \) and \( T \), expression (A.2) for the Helmholtz free energy is extended as

\[ dA = -SdT - PdV + \mu dN, \quad (A.10) \]

where \( \mu \) is called the chemical potential,

\[ \mu = \left( \frac{\partial A}{\partial N} \right)_{T,V}. \quad (A.11) \]
Bibliography


[44] Stell G., in Ref. [18], pg. II-171.


Index

Algorithm
Verlet, 45

Bond, 53

Canonical partition function, 14
Closure, 62
Computer experiment, 41
Configurational integral, 16

De Broglie wavelength, 15, 16

Diagram
irreducible, 57
reducible, 57

Diagrammatic technique, 50

Equation
of state, compressibility, 35
Ornstein-Zernike, 62
viral, of state, 33

Field point, 53

Fluids
dipolar, 20
molecular, 20
normal, 20
polar, 20
quadrupolar, 20
simple, 20

Function
background, correlation, 37
total, correlation, 36
correlation, 31, 32, 35
density-density, correlation, 36
distribution, 31
Mayer, 19
site-site, correlation, 39

Functional differentiation, 50

Gear integrator, 46

Helmholtz free energy, 14

Hydrogen bonding, 19, 20

Method
Euler, 44
predictor-corrector, 44

Monte Carlo
naive, 48
smart, 49

Node, 57

Pairwise additivity, 18

Parameter
scaling, 33

Perturbation, 65

Potential
Buckingham, 23
Buckingham, modified, 23
effective pair, 18
hard sphere, 26
Lennard-Jones, 24
m-6, 24
of the mean force, 37
reference, 65
INDEX

Yukawa, hard-core, 25

Reference point, 14, 18
Running coordination number, 37

Simulation
  computer, 41
  molecular, 42
Site-site model, 27
Square-well, 26

Topological reduction, 50