SIMULATING THE REAL WORLD

A hierarchy of models from quantum mechanics to fluid dynamics

Herman J. C. Berendsen
emeritus professor of physical chemistry,
the University of Groningen, the Netherlands

1

INTRODUCTION

1.1 What is this book about?

1.1.1 Simulation of real systems

Computer simulations of real systems require a *model* of that reality. A model consists of both *a representation of the system* and *a set of rules* that describe the behaviour of the system. For dynamical descriptions one needs in addition a specification of the *initial state* of the system, and if the response to external influences is required, a specification of the *external influences.*

Both the model and the method of solution depend on the *purpose* of the simulation: they should be *accurate* and *efficient.* The model should be chosen accordingly. For example: an accurate quantum-mechanical description of the behaviour of a many-particle system is not efficient for studying the flow of air around a moving wing; on the other hand, the Navier-Stokes equations – efficient for fluid motion – cannot give an accurate description of the chemical reaction in an explosion motor. Accurate means that the simulation will reliably (within a required accuracy) predict the real behaviour of the real system, and efficient means ‘feasible with the available technical means’. This combination of requirements rules out a number of questions; whether a question is answerable by simulation depends on

- the state of theoretical development (models and methods of solution),
- the computational capabilities,
- the possibilities to implement the methods of solution in algorithms,
- the possibilities to validate the model.

Validation means the assessment of the accuracy of the model (compared to physical reality) by critical experimental tests. Validation is a crucial part of modelling.

1.1.2 System limitation

We limit ourselves to models for the real world around us. This is the realm of chemistry, biology and material sciences, and includes all industrial and practical applications. We do not include the formation of stars and galaxies (*stellar dynamics*) or the physical processes in hot plasma on the sun's surface (*astrophysics*); neither do we include the properties and interactions of elementary particles (*quantum chromodynamics*) or processes in atomic nuclei or neutron stars. And, except for the purposes of validation and demonstration, we shall not consider unrealistic models that are only meant to test a theory. Summarizing: we shall look at literally 'down-to-earth' systems consisting of atoms and molecules under non-extreme conditions of pressure and temperature.

This limits our discussion in practice to systems that are made up of interacting *atomic nuclei* which are specified by their mass, charge and spin, *electrons*, and *photons* that carry the electromagnetic interactions between the nuclei and electrons. Occasionally we may wish to add gravitational interactions to the electromagnetic ones. The internal structure of atomic nuclei is of no consequence for the behaviour of atoms and molecules, (if we disregard radioactive decay): nuclei are so small with respect to the spatial spread of electrons that only their *monopole* properties as total charge and total mass are important. Nuclear excited states are so high in energy that they are not populated at reasonable temperatures. Only the spin degeneracy of the nuclear ground state plays a role when nuclear magnetic resonance is considered; in that case the nuclear magnetic dipole and electric quadrupole moment are important as well.

In the normal range of temperatures this limitation implies a practical division between electrons on the one hand and nuclei on the other: while all particles obey the rules of quantum mechanics, the quantum character of electrons is essential but the behaviour of nuclei approaches the classical limit. This distinction has far-reaching consequences, but it is rough and inaccurate. For example, protons are light enough to violate the classical rules. The validity of the classical limit will be discussed in detail in this book.

1.1.3 Sophistication versus brute force

Our interest in *real* systems rather than simplified model systems is consequential for the kind of methods that can be used. Most real systems concern some kind of condensed phase: they (almost) never consist of isolated molecules and can (almost) never be simplified because of inherent symmetry. Interactions between particles can (almost) never be described by mathematically simple forms and often require numerical or tabulated descriptions. Realistic systems usually consist of a very large number of interacting particles, embedded in some kind of environment. Their behaviour is (almost) always determined by statistical averages over ensembles consisting of elements with random character, as the random distribution of thermal kinetic energy over the available degrees of freedom. That is why statistical mechanics plays a crucial role in this book.

The complexity of real systems prescribes the use of methods that are easily extendable to large systems with many degrees of freedom. Physical theories that apply to simple models only, will (almost) always be useless. Good examples are the very sophisticated statistical-mechanical theories for atomic and molecular fluids, relating fluid structural and dynamic behaviour to interatomic interactions. Such theories work for atomic fluids with simplified interactions, but become inaccurate and intractable for fluids of polyatomic molecules or for interactions that have a complex form. While such theories thrived in the 1950's to 1970's, they have been superseded by accurate simulation methods, which are faster and easier to understand, while they predict liquid properties from interatomic interactions much more accurately. Thus sophistication has been superseded by brute force, much to the dismay of the sincere basic scientist.

Many mathematical tricks that employ the simplicity of a toy model system cannot be used for large systems with realistic properties. In the box below an example is given of the brute-force approach to a problem that has a simple and elegant solution. To apply such a brute-force method to a simple problem seems outrageous and intellectually very dissatisfying. Nevertheless, the elegant solution cannot be readily extended to many particles or complicated interactions, while the brute-force method can. Thus not only sophistication in physics, but also in mathematics, is often replaced by brute force methods. There is an understandable resistance against this trend among well-trained mathematicians and physicists, while scientists with a less elaborate training in mathematics and physics welcome the opportunity to study complex systems in their field of application. The field of simulation has made theory much more widely applicable and has become accessible to a much wider range of scientists than before the 'computer age'. Simulation has become a 'third way' of doing science, not in stead of, but in addition to theory and experimentation.

There is a danger, however, that applied scientists use 'standard' simulation methods, even worse use 'black-box' software, without realizing on what assumptions the methods rest and what approximations are implied.

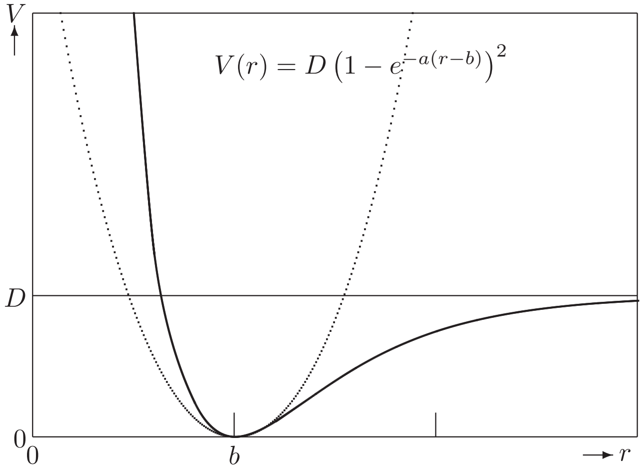


Fig. 1.1. Morse curve with $a = 2/b$ (drawn line). Dotted curve: parabola with same curvature as Morse curve at $r = b$: $V = Da^2(r - b)^2$.

This book is meant to provide the necessary scientific background and to promote awareness for the limitations and inaccuracies of simulating the 'real world'.

Example 1.1 An oscillating bond

In this example we consider the classical bond length oscillation of a simple diatomic molecule, using the molecule hydrogen fluoride (HF) as an example. In the simplest approximation the potential function is a parabola:

$$V(r) = \frac{1}{2}k(r - b)^2, \tag{1.1}$$

with r the H-F distance, k the force constant and b the equilibrium distance. A better description of the potential function is the *Morse function* (see fig. 1.1)

$$V(r) = D\left(1 - e^{-a(r-b)}\right)^2, \tag{1.2}$$

where D is the dissociation energy and a is a constant related to the steepness of the potential. The Morse curve is approximated near the minimum at $r = b$ by a parabola with force constant $k = 2Da^2$.

The Morse curve is only a convenient analytical expression that has some essential features of a diatomic potential, but there is no theoretical justification for this particular form. In many occasions we may not even have an analytical form for the potential, but know the potential at a number of discrete points, e.g. from quantum-chemical calculations. In that case the best way to proceed is to construct the potential function from *cubic spline interpolation* of the computed points. Because

Fig. 1.2. Oscillation of the HF bond length, simulated with (a) the harmonic oscillator, (b) the Morse curve with initial deviation from the equilibrium bond length of -0.01 nm, (c) *idem* with initial deviation of 0.05 nm, showing increased anharmonic behaviour.

cubic splines (see TOPIC **??**) have continuous second derivatives, the forces will behave smoothly as they will have continuous first derivatives everywhere.

A little elementary mechanics shows that we can split off the translational motion of the molecule as a whole, and that - in the absence of rotational motion - the bond will vibrate according to the equation of motion:

$$\mu \ddot{r} = -\frac{dV}{dr} \tag{1.3}$$

where $\mu = m_H m_F/(m_H + m_F)$ is the reduced mass of the two particles. When we start at time $t = 0$ with a displacement Δr and zero velocity, the solution for the harmonic oscillator is

$$r(t) = b + \Delta r \cos \omega t, \tag{1.4}$$

with $\omega = \sqrt{k/\mu}$. So the analytical solution is simple, and we do not need any numerical simulation to derive the frequency of the oscillator. For the Morse oscillator the solution is not as straightforward, although we can predict that it should look much like the harmonic oscillator with $k = 2Da^2$ for small-amplitude vibrations. But we may expect anharmonic behaviour for larger amplitudes. Now numerical simulation is the easiest way to derive the dynamics of the oscillator. For a spline-fitted potential we *must* resort to numerical solutions. The extension to more complex problems, like the vibrations of a molecule consisting of several interconnected harmonic oscillators, is quite straightforward in a simulation program, while analytical solutions require sophisticated mathematical techniques.

The reader is invited to write a simple molecular dynamics program that uses the following routine `mdstep` to perform one MD step with the *velocity-Verlet* algorithm (see **??**), define a function `force` that provides an array of forces $\boldsymbol{F}$, given the coordinates $\boldsymbol{r}$, both for the harmonic and the Morse potential, and try out a few initial conditions and time steps. As a rule of thumb: start with a time step such that the fastest oscillation period contains 50 steps. You may generate curves like those in fig. 1.2. See what happens if you give the molecule a rotational velocity! Keep to 'molecular units': *mass*: u, *length*: nm, *time*: ps, *energy*: kJ/mol. Use the following data for hydrogen fluoride:

mass H	m_{H}	1.0079	u
mass F	m_{F}	18.9984	u
dissoc. const.	D	569.87	kJ/mol
equil. bond length	b	0.09169	nm
force constant	k	5.82×10^5	$\mathrm{kJ\,mol^{-1}\,nm^{-2}}$

The following function performs one velocity-Verlet time step of MD on a 3-dimensional molecule. Given masses m (as list), initial positions r, velocities v *and forces F (as n*3 arrays), it returns r, v and F one timestep later. A function* `force(r)` *must be provided that returns F, given r. A workspace Ftemp (n*3) must be globally defined. The function will modify r,v,F in place.*

```
def mdstep(m,r,v,F,force,timestep):
    n=len(m)
    for i in range(n):                                      update r
        temp=0.5*timestep/m[i]
        for j in range(3):
            Fij=F[i][j]
            r[i][j]=r[i][j]+(v[i][j]+temp*Fij)*timestep
            Ftemp[i][j]=Fij                           store old force
    F=force(r)                                compute force at new r
    for i in range(n):                                      update v
        temp=0.5*timestep/m[i]
        for j in range(3):
            v[i][j]=v[i][j]+temp*(F[i][j]+F1[i][j])
    return r,v,F
```

1.2 A modelling hierarchy

The behaviour of a system of particles is in principle described by the rules of relativistic quantum mechanics. This is – within the limitation of our system choices – the highest level of description. We shall call this *level 1.* All other levels of description, such as considering atoms and molecules in stead of nuclei and electrons, classical dynamics in stead of quantum dynamics, or continuous media in stead of systems of particles, represent approximations to level 1. These approximations can be ordered in a hierarchical sense from fine atomic detail to course macroscopic behaviour. Every lower level looses detail and looses applicability or accuracy for a certain class of systems and questions, but gains applicability or efficiency for another class of systems and questions. The following scheme lists several levels in this hierarchy.

LEVEL 1 RELATIVISTIC QUANTUM DYNAMICS

System	**Rules**
Atomic nuclei (mass, charge, spin), electrons (mass, charge, spin), photons (frequency)	Relativistic time-dependent quantum mechanics; Dirac's equation; (quantum) electrodynamics

Approximation
particle velocities small compared to velocity of light

No Go
electrons close to heavy nuclei; hot plasma's

LEVEL 2 QUANTUM DYNAMICS

System	**Rules**
Atomic nuclei, electrons, photons	Non-relativistic time-dependent Schrödinger eqn (TDSE); time-independent Schrödinger eqn; Maxwell eqns

Approximation
Born-Oppenheimer approx.: electrons much faster than nuclei

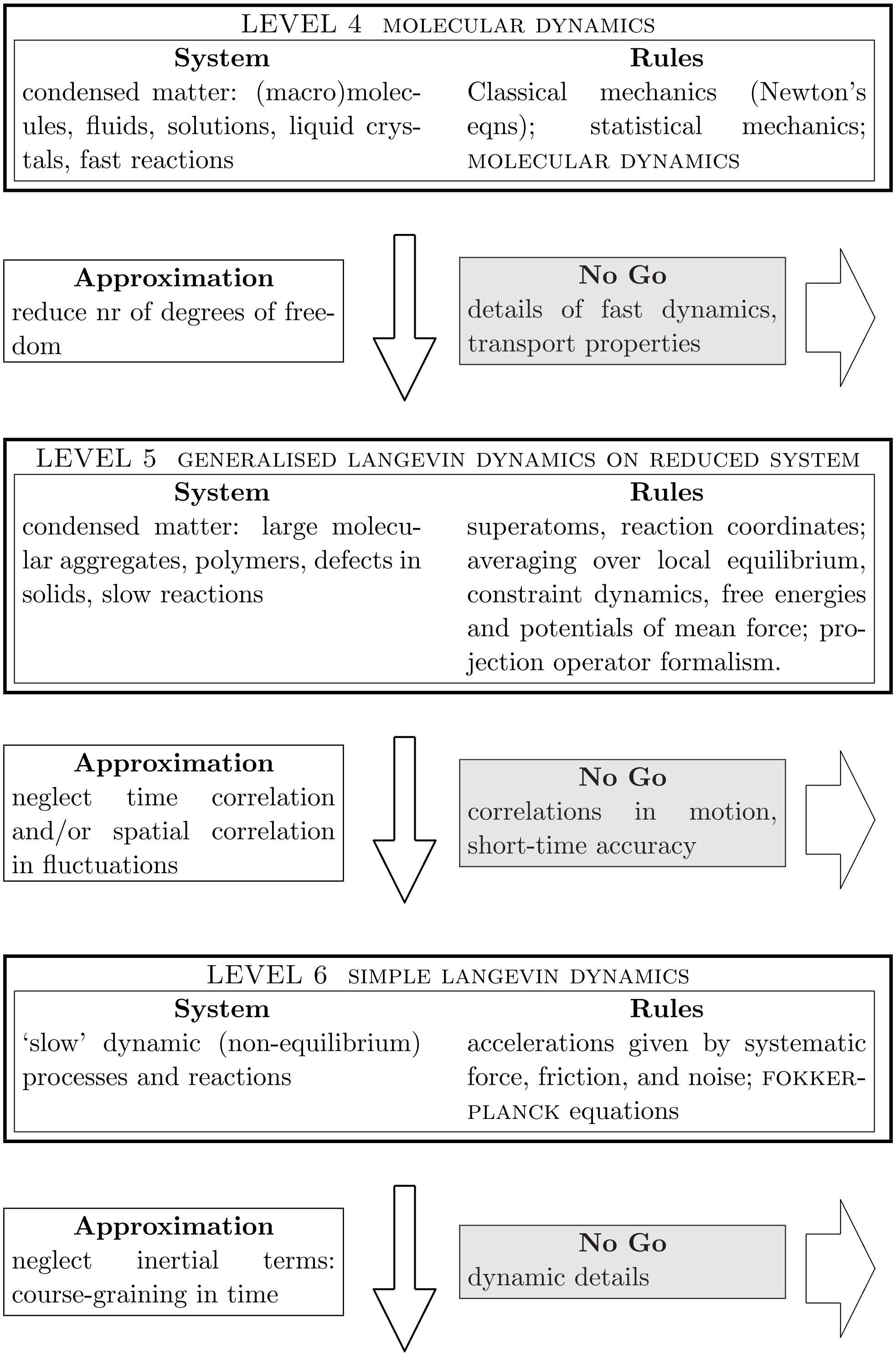
No Go
electron dynamics (e.g. in semiconductors); fast electron transfer processes; dynamic behaviour of excited states

LEVEL 3 ATOMIC QUANTUM DYNAMICS

System	**Rules**
atoms, ions, molecules, (photons)	atoms move in effective potential due to electrons; atoms may behave according to TDSE

Approximation
atomic motion is classical

No Go
proton transfer; hydrogen and helium at low temperatures; fast reactions and high-frequency motions



LEVEL 4 MOLECULAR DYNAMICS

System	Rules
condensed matter: (macro)molecules, fluids, solutions, liquid crystals, fast reactions	Classical mechanics (Newton's eqns); statistical mechanics; MOLECULAR DYNAMICS

Approximation
reduce nr of degrees of freedom

No Go
details of fast dynamics, transport properties

LEVEL 5 GENERALISED LANGEVIN DYNAMICS ON REDUCED SYSTEM

System	Rules
condensed matter: large molecular aggregates, polymers, defects in solids, slow reactions	superatoms, reaction coordinates; averaging over local equilibrium, constraint dynamics, free energies and potentials of mean force; projection operator formalism.

Approximation
neglect time correlation and/or spatial correlation in fluctuations

No Go
correlations in motion, short-time accuracy

LEVEL 6 SIMPLE LANGEVIN DYNAMICS

System	Rules
'slow' dynamic (non-equilibrium) processes and reactions	accelerations given by systematic force, friction, and noise; FOKKER-PLANCK equations

Approximation
neglect inertial terms: course-graining in time

No Go
dynamic details

LEVEL 7 BROWNIAN DYNAMICS

System	**Rules**
coarse-grained non-equilibrium processes; colloidal systems; polymer systems	velocities given by force and friction, plus noise; Brownian (diffusive) dynamics; Onsager flux/force relations

Approximation
reduce description to continuous densities of constituent species

No Go
details of particles

LEVEL 8 MESOSCOPIC DYNAMICS

System	**Rules**
as for level 7; self-organizing systems; reactive nonequilibrium systems	density description: mass conservation plus dynamic flux equation, with noise.

Approximation
average over ‘infinite’ number of particles

No Go
spontaneous structure formation driven by fluctuations

LEVEL 9 REACTIVE FLUID DYNAMICS

System	**Rules**
non-equilibrium macroscopic mixture of different species (as the atmosphere for weather forecasting	energy, momentum and mass conservation; reactive fluxes

Approximation
reduce to one species with Newtonian viscosity

No Go
reactive processes; non-Newtonian behaviour

LEVEL 10 FLUID DYNAMICS

System	**Rules**
non-equilibrium macroscopic fluids: gases and liquids	energy, momentum and mass conservation; NAVIER-STOKES equations

Approximation
low fluid velocities (low Reynolds number)

No Go
turbulence

LEVEL 11 STEADY-FLOW FLUID DYNAMICS

System	**Rules**
non-equilibrium fluids with laminar flow	simplified Navier-Stokes equation

From level 5 onward, not all atomic details are included in the system description: one speaks of *coarse-graining in space*. From level 6 onward dynamic details on a short time scale are disregarded by *coarse-graining in time*.

In the last stages of this hierarchy (levels 8 to 11), the systems are not modelled by a set of particles, but rather by properties of a *continuum*. Equations describe the time evolution of the continuum properties. Usually such equations are solved on a spatial grid using finite difference or finite elements methods for discretizing the continuum equations. A different approach is the use of particles to represent the continuum equations, called *dissipative particle dynamics* (DPD). The particles are given the proper interactions representing the correct physical properties that figure as parameters in the continuum equations.

Note that we have considered dynamical properties at all levels. Not all questions we endeavor to answer involve dynamic aspects, such as the prediction of static equilibrium properties (e.g. the binding constant of a ligand to a macromolecule or a solid surface). For such static questions the answers may be found by sampling methods, such as Monte Carlo simulations, that generate a representative statistical ensemble of system configurations rather than a trajectory in time.

1.3 Trajectories and distributions

Dynamic simulations of many-particle systems contain fluctuations or stochastic elements, either due to the irrelevant particular choice of initial conditions (as the exact initial positions and velocities of particles in a classical simulation or the specification of the initial wave function in a quantum-dynamical simulation), or due to the 'noise' added in the method of solution (as in Langevin dynamics where a stochastic force is added to replace forces due to degrees of freedom that are not explicitly represented). Fluctuations are implicit in the dynamic models up to and including LEVEL 8.

The precise details of a particular trajectory of the particles have no relevance for the problem we wish to solve. What we need is always an average over many trajectories, or at least an average property, such as the average or the variance of a single observable or a correlation function, over one long trajectory. In fact, an individual trajectory may even have *chaotic* properties: two trajectories with slightly different initial conditions may deviate drastically after a sufficiently long time. However, the average behaviour is deterministic for most physical systems of interest.

In stead of generating distribution functions and correlation functions from trajectories, we can also try to define equations, such as the Fokker-Planck equation, for the distribution functions (probability densities) or correlation functions themselves. Often the latter is very much more difficult than generating the distribution functions from particular trajectories. An exception is the generation of equilibrium distributions, for which Monte Carlo methods are available that circumvent the necessity to solve specific equations for the distribution functions. Thus the simulation of trajectories is often the most efficient — if not the only possible — way to generate the desired average properties.

While the notion of a trajectory as the time evolution of positions and velocities of all particles in the system, is quite valid and clear in classical mechanics, there is no such notion in quantum mechanics. The description of a system in terms of a wave function Ψ is by itself a description in terms of a *probability density*: $\Psi^*\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n, t)$ is the probability density that the particles $1, \ldots, n$ are at positions $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n$ at time t. Even if the initial state is precisely defined by a sharp wave function, the wave function evolves under the quantum-dynamical equations to yield a probability distribution rather than a precise trajectory. From the wave function evolution expectation values (i.e., average properties over a probability distribution) of physical observables can be obtained by the laws of quantum mechanics, but the wave function cannot be interpreted as the (unmeasurable) property of a single particle.

Such a description fits in well with equations for the evolution of probability distributions in classical systems, but it is not compatible with descriptions in terms of classical trajectories. This fundamental difference in interpretation lies at the basis of the difficulties we encounter if we attempt to use a hybrid quantum/classical description of a complex system. If we insist on a trajectory description, the quantum-dynamical description should be reformulated by some kind of contraction and sampling to yield trajectories that have the same statistical properties as prescribed by the quantum evolution. It is for the same reason of incompatibility of quantum descriptions and trajectories that quantum corrections to classical trajectories cannot be unequivocally defined, while quantum corrections to equilibrium probability distributions can be systematically derived.

1.4 Books along the way

1 S. Gasiorowicz, *Quantum Physics* (1974) is a readable text book on quantum physics with a discussion of the limits of classical physics.

2 L.I. Schiff, *Quantum Mechanics* (1968), a compact classic text book, slightly above the level of Gasiorowicz.

3 E. Merzbacher, *Quantum Mechanics* (1970) is another classic text book with a complete coverage of the main topics.

4 L.D. Landau ands E.M. Lifshitz, *Quantum Mechanics (Non-relativistic Theory)* (1987): This is one volume in the excellent series "Course of Theoretical Physics". Its level is advanced and sophisticated.

5 P.A.M. Dirac, *The Principles of Quantum Mechanics* (1958) by one of the founders of quantum mechanics: advisable reading only for the dedicated student.

6 F.S. Levin, *An Introduction to Quantum Theory* (2002) introduces principles and methods of basic quantum physics at great length. It has a part on 'complex systems' that does not go far beyond two-electron atoms.

7 A. Szabo and N.S. Ostlund, *Modern Quantum Chemistry* (1982) is a rather complete and reasonably modern text book on quantum chemistry, entirely devoted to the solution of the time-independent Schrödinger equation for molecules.

8 R. McWeeny, *Methods of Molecular Quantum Mechanics* (1992) is the classical text on quantum chemistry.

9 R.G. Parr and W. Yang *Density Functional Theory* (1989). An early, and one of the few books on the still-developing DFT.

10 F. Jensen, *Introduction to Computational Chemistry* (1999). A modern comprehensive survey of methods in computational chemistry including a range of ab initio and semi-empirical quantum chemistry methods, but also molecular mechanics and dynamics.

11 H. Goldstein, *Classical Mechanics* (1980) is the classical text and reference book on mechanics. The revised third edition (Goldstein *et al.*, 2002) has an additional chapter on chaos, as well as other extensions, at the expense of details that were present in the first two editions.

12 L.D. Landau and E.M. Lifshitz, *Mechanics* (1976). Not as complete as Goldstein, but superb in its development of the theory.

13 L.D. Landau and E.M. Lifshitz, *Statistical Physics* (1976). Basic text for statistical mechanics.

14 K. Huang, *Statistical Mechanics* (1963). Statistical mechanics text book from a physical point of view, written before the age of computer simulation.

15 T.L. Hill, *Statistical Mechanics* (1956): A classic and complete, but now somewhat outdated, statistical mechanics text book with due attention to chemical applications. Written before the age of computer simulation.

16 D.A. McQuarrie, *Statistical Mechanics* (1973) is a high quality text book, covering both physical and chemical applications.

17 M. Toda, R. Kubo and N. Saito, *Statistical Physics. I. Equilibrium statistical mechanics* (1983) and R. Kubo, M. Toda and N. Hashitsume *Statistical Physics. II. Nonequilibrium statistical mechanics* (19835) emphasize physical principles and applications. These texts were originally published in Japanese in 1978. Especially vol. II is a good reference for linear response theory, both quantum-mechanical and classical, to which Kubo has contributed significantly. Not recommended for chemists.

18 D. Chandler, *Introduction to Modern Statistical Mechanics* (1987): A basic statistical mechanics text book emphasizing fluids, phase transitions and reactions, written in the age of computer simulations.

19 B. Widom, *Statistical Mechanics, A concise introduction for chemists* (2002) is what it says: an introduction for chemists. It is well-written, but does not reach the level to treat the wonderful inventions in computer simulations, such as particle insertion methods, for which the author is famous.

20 M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids* (1987). A practical guide to molecular dynamics simulations with emphasis on the methods of solution rather than the basic underlying theory.

21 D. Frenkel and B. Smit, *Understanding Molecular Simulation* (1996). A modern, instructive, and readable book on the principles and practise of Monte Carlo and Molecular Dynamics simulations. The scope is somewhat limited because the authors insist on precision.

22 A.R. Leach, *Molecular Modelling, Principles and Applications* (1996) aims at the simulation of molecular systems leading up to drug discovery. Starting with quantum chemistry, the book decribes energy minimization, molecular dynamics and Monte Carlo methods in detail.

23 N.G. van Kampen *Stochastic Processes in Physics and Chemistry* (1981) gives a very precise and critical account of the use of stochastic and Fokker-Planck type equations in (mostly) physics and (a bit) chemistry.

24 R. Kubo, M. Toda and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics* (1985) by one of the pioneers (Kubo) of the statistical-mechanical linear response theory. Describes the connection between correlation functions and macroscopic relaxation.

25 H. Risken *The Fokker-Planck equation* (1989) treats the evolution of probability densities.

26 C.W. Gardiner *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (1985) is a reference book for modern developments in stochastic dynamics. It treats the relations between stochastic equations and Fokker-Planck equations.

27 M. Doi and S.F. Edwards, *The Theory of Polymer Dynamics* (1986) is the already classic introduction to mesoscopic treatment of polymers.

28 L.D. Landau and E.M. Lifshitz, *Fluid Mechanics* (1987) is an excellent account of the physics behind the equations of fluid dynamics.

29 T. Pang *Computational Physics* (1997). A modern and versatile treatise on methods in computational physics, covering a wide range of applications. The emphasis is on the computational aspects of the methods of solution, not on the physics behind the models.

30 F.J. Vesely, *Computational Physics, An Introduction* (2nd ed., 2001) is an easily digestable treatment of computational problems in physics, with emphasis on mathematical and computational methods rather than on the physics behind the equations.

31 M. Griebel, S. Knapek, G. Zumbusch and A. Caglar, *Numerische*

Simulation in der Moleküldynamik (2003) gives many advanced details on methods and algorithms for dynamic simulation with particles. The emphasis is on computational methods including parallelization techniques; programs in C are included. Sorry for some readers: the text is in German.