

SOMETHING SOMETHING COMPUTATIONAL  
QUANTUM CHEMISTRY SOMETHING SOMETHING  
NEURAL NETWORKS

by

Morten Ledum

THESIS

for the degree of

MASTER OF SCIENCE



Faculty of Mathematics and Natural Sciences  
University of Oslo

December 2017



# Abstract

I havent done much worthwhile but hopefully I can bullshit my way into not failing.



# Acknowledgements

I would like to thank my supervisors (in no particular order) Anders Malthe-Sørensen, Morten Hjorth-Jensen, and Anders Håfreager for their support during the years as a master student here at the physics department. Along with the rest of the community at the computational physics group, you have created a wonderful environment for learning that I will sorely miss. To Morten: thank you for taking me under your wing here at the computational physics group. It has been a pleasure to be your student and later your TA. I am proud to have taken some small part in teaching the introductory computational physics course with you these past several years.

To my tireless *room-mate* here in office V306, Håkon Kristiansen. It feels like we have been sitting here for decades, and I very much appreciate the company. I promise to stop asking you difficult questions you cant answer (at some point). Maybe one day the dream will come true and someone will pay us to just sit around and *figure out* whatever is interesting to us that day.

I want to acknowledge the two people who have probably influenced me as a student the most. Anders Håfreager and Håvard Tveit Ihle, you both are an inspiration to me, showing what you can achieve through diligent hard work. Anders, thank you for teaching me how to write (less awful) code, and for helping me not fail QFT. I truly appreciate the endless enthusiasm and the willingness to go out of your way to help regardless of the nature of the problem. Håvard, I am looking forward to floundering my way through your cosmology lectures this spring!

I also wish to thank the unlikely Johan, the (true[ly]) interesting Icelandic bat-camel for all the stimulating distractions throughout the time working on this thesis.

Lastly, to Vilde: Thank you for the endless encouragement (and all the sushi!). Thank you for putting up with the largely absent and massively stressed out version of myself for long months.

Morten Ledum  
Oslo, December 2017



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Quantum and classical dynamics . . . . .	2
1.2	Neural networks in molecular dynamics . . . . .	2
1.3	Goals . . . . .	3
1.4	Structure of the thesis . . . . .	3
<b>I</b>	<b>Foundational theory</b>	<b>5</b>
<b>II</b>	<b>Conclusions and future work</b>	<b>7</b>
<b>Appendices</b>		
<b>Appendix A</b>	<b>Natural units: Hartree atomic units</b>	<b>11</b>
<b>Appendix B</b>	<b>Basics of numerical integration</b>	<b>13</b>
<b>Appendix C</b>	<b>Functionals and functional variations</b>	<b>19</b>
C.0.1	Short mathematical interlude . . . . .	20
C.0.2	Functional differentials and derivatives . . . . .	20





# Chapter 1

## Introduction

After being developed and pioneered in conjunction with nuclear physics during World War II, computer modelling and numerical experiments have become ubiquitous in the natural sciences. The list of areas in which simulations have been used to produce significant new results encompass now pretty much all of them [citation needed]. Today computer simulations play as natural a part of the hard sciences as laboratory experiments and theory. In most cases all three are used in order to glean new scientific insight. ((examples here?))

The advent of computer simulations during the last several decades have in particular made it possible to study moderately sized quantum mechanical systems from first principles. As our ability to solve—in closed form—the governing equations of quantum mechanics (QM) vanishes extremely quickly as the number of constituent particles exceed just a few, numerics are used to augment the proverbial *pen and paper*. It is striking that the underlying theory for all of chemistry and most of physics have been known for almost a century but the problem preventing us from essentially *solving* chemistry is almost purely computational: the equations resulting from the exact application of this theoretical framework are way too difficult to solve.

Any approximative scheme which aims to solve the many-body Schrödinger equation from scratch subject to some (more or less) well-defined simplifications is called an *ab initio* method. Working from first principles the aim of such algorithms extract information from a theoretical QM system in a reasonable amount of time. In order to accomplish this, a number of complicating intricacies need to be disregarded. The magnitude of the simplifications—essentially the number and importance of complicating factors dropped—determine both the efficacy and the efficiency of the method: More simplifications made allow solutions to be found for larger systems (albeit less precise solutions), whereas extremely precise solutions can be found for small systems if very few simplifications are employed.

Despite tremendous increases in available numerical computational power in the latter half of the previous-, and the early parts of the current century, any such approximate scheme used is still heavily limited w.r.t. the system size. In practice, most methods are limited to systems of containing on the order of between  $10^2$  (for high-

precision methods such as CI, CC, DMC, etc.) and  $10^5$  electrons (for faster HF and DFT methods) [1, 2, 3]. Extracting information from larger systems necessitate the use of classical or semi-classical algorithms, such as molecular dynamics (MD). Using MD, the time evolution of up to around  $10^7$  particles can feasibly be simulated over the order of nano seconds [4, 5]. Corresponding large-scale cosmological  $N$ -body simulations have been run for as many as  $10^{11}$  particles [6, 7]. As macroscopic objects contain on the order of  $10^{23}$  atoms, we are ostensibly a long way away from even classically modelling the constituent particles directly. In this region, continuum models are used.

## 1.1 Quantum and classical dynamics

As previously noted, solving the Schrödinger equation (SE) exactly by hand is impossible in the overwhelming majority of interesting cases. However, methods which can get close to the exact solution exists. Full Configuration Interaction (FCI) or direct diagonalization of the Hamiltonian is exact in the limit of an infinite orbital basis set but suffers from an exponential complexity scaling (in system *and* basis size) [8]. The related Configuration Interaction (CI) and Coupled Cluster (CC) approaches both truncate the FCI expansion of Slater determinants, thus gaining speed but loosing some accuracy [9, 10]. Diffusion Monte Carlo (DMC) techniques can in principle provide the exact solution to the SE by imaginary-time evolution of an initial wave function guess [11, 12]. In practice, DMC methods are highly dependent on this ansatz and thus require as input the results of less accurate method but faster methods. One example may be the Variational Monte Carlo (VMC) method: conceptually simpler and faster than DMC, but not as accurate [12, 13, 14].

The Hartree-Fock (HF) framework—which provides an efficient but not enormously accurate result—has seen extensive use since its inception in 1930 [15, 16, 17]. However, by far the most popular approximation is Density Functional Theory (DFT), developed by W. Kohn and L. J. Sham in 1965 [18, 19]. Between 1980 and 2010, DFT was the most active field in physics with eight out of the top ten most cited papers being on the subject [20].

"This is where multiscale modeling comes in. By considering simultaneously models at different scales, we hope to arrive at an approach that shares the efficiency of the macroscopic models as well as the accuracy of the microscopic models." Weinan: Principles of Multiscale modelling side viii

## 1.2 Neural networks in molecular dynamics

Neural networks can parameterize the data from millions of electronic structure calculations and in theory provide accuracy close to the direct quantum mechanical calculations at a fraction of the calculational cost.

## 1.3 Goals

The goals of the thesis are as follows:

## 1.4 Structure of the thesis

This thesis is split into four parts. The first part, *foundational theory* presents an overview of classical mechanics (very briefly) and quantum mechanics as relevant for molecular dynamics and electronic structure calculations. In addition, the basics of machine learning and artificial neural networks are presented. Part one lays the groundwork and establishes most of the notation used later in part two: *Advanced theory*. Here, the different approximative schemes for solving the quantum equations of motion are presented. Alongside them is the theory of molecular dynamics simulations and advanced topics in neural networks.

The penultimate part contains information w.r.t. the concrete implementation of frameworks described in part two. Key parts of each code base are outlined in detail. Also contained in part three are validation tests of each implementation. Lastly, results for the full work-flow utilizing all the developed programs are presented.

The thesis ends with part four, containing conclusions and prospect for future work.



# **Part I**

## **Foundational theory**



## **Part II**

### **Conclusions and future work**





# **Appendices**



# Appendix A

## Natural units: Hartree atomic units

When working within a specific branch of physics, it is often useful to deviate from the every-day SI units of measurements and instead use units which are *natural* to the systems under study. Since we are working with "small" systems, the SI *meter*, *second*, *kilogram*, and *coulomb* are of little use to us. Instead we will work in a system of units in which we define the mass of the electron,  $m_e$ , to be the scale by which we measure all other masses. This obviously means the numerical value of the electron mass becomes unity,  $m_e = 1$ . In the same way, we will use Planck's constant,  $\hbar$ , as the scale by which we measure angular momentum and action, the electron charge,  $e$ , will be our scale for electrical charge, and finally Coulomb's constant,  $k_e$ , will be our scale of electric permittivity.

The usual way to state this is to set  $\hbar = e = m_e = k_e = 1$ , and the system of units derived from these four definitions is called Hartree atomic units. We can think of this as the *natural* system of units for the Hydrogen atom system. To better see why this is the case, let us combine these four quantities in such a way as to produce a length.

In terms of the four fundamental dimensions of physics: Length(L), time(T), mass(M), and charge(C), the units of  $\hbar$ ,  $m_e$ ,  $e$ , and  $k_e$  are  $[\hbar] = \text{ML}^2\text{T}^{-1}$ ,  $[m_e] = \text{M}$ ,  $[e] = \text{C}$ , and  $[k_e] = \text{ML}^3\text{C}^{-2}\text{T}^{-2}$ , respectively. Combining arbitrary powers of these four constants gives

$$\begin{aligned} [\lambda(a, b, c, d)] &= [k_e^a \hbar^b m_e^c e^d] = (\text{M}^a \text{L}^{3a} \text{C}^{-2a} \text{T}^{-2a}) (\text{M}^b \text{L}^{2b} \text{T}^{-b}) (\text{M}^c) (\text{C}^d) \\ &= \text{L}^{2a+3b} \text{T}^{-a-2b} \text{M}^{a+b+c} \text{C}^{-2b+d}. \end{aligned} \quad (\text{A.1})$$

There is exactly one way to realize a length from these exponents, i.e. solving the four equations  $2a + 3b = 1$ ,  $-a - 2b = 0$ ,  $a + b + c = 0$ , and  $-2b + d = 0$ :  $a = -1$ ,  $b = 2$ ,  $c = -1$ , and  $d = -2$ . This means that the natural length scale of our problem is simply (up to a numerical constant)

$$\text{L}_{\text{scale}} = a_0 = k_e^{-1} \hbar^2 m_e^{-1} e^{-2} = \frac{\hbar^2}{k_e m_e e^2} = \frac{4\pi\epsilon_0 \hbar^2}{m_e e^2}, \quad (\text{A.2})$$

which we recognize as simply the *Bohr radius*.

We can go through this same exercise to find a natural *time* scale for our system. There is a unique way to combine the exponents  $a$ ,  $b$ ,  $c$ , and  $d$  in order to realize a time, namely  $a = -2$ ,  $b = 3$ ,  $c = -1$ ,  $d = -4$ , or

$$T_{\text{scale}} = k_e^{-2} \hbar^3 m_e^{-1} e^{-4} = \frac{\hbar^3}{k_e^2 m_e e^4} = \frac{\hbar a_0}{k_e e^2}. \quad (\text{A.3})$$

This is the revolution time of an electron in the lowest lying hydrogen state in the Bohr model (apart from a factor of  $2\pi$ ).

From  $a_0$  and  $T_{\text{scale}}$  we can find the natural energy scale,

$$E_{\text{scale}} = m_e \frac{a_0^2}{a_0^2 \left( \frac{\hbar}{k_e e^2} \right)^2} = \frac{m_e k_e^2 e^4}{\hbar^2} \equiv E_h, \quad (\text{A.4})$$

which we will call a Hartree.

Finally, before we go on we may use the expression for the *fine structure constant* to find the numerical value of  $c$  in this system. From

$$\alpha = \frac{k_e e^2}{\hbar c} \Rightarrow c = \frac{k_e e^2}{\hbar \alpha} = \frac{1}{\alpha} \simeq 137, \quad (\text{A.5})$$

after substituting  $\hbar = e = k_e = 1$ .

# Appendix B

## Basics of numerical integration

### Riemann integral and Riemann integrable functions

Given a function  $f(x)$  and a closed finite subset of  $\mathbb{R}$ ,  $[a, b]$  with  $a < b$ , a *Riemann sum* of  $f$  is defined as the sum of values attained on  $n$  sub-intervals of  $[a, b]$ , i.e.

$$S_n = \sum_{i=1}^n (x_i - x_{i-1}) f_i. \quad (\text{B.1})$$

The  $x_i$ s here define the partitioning into sub-intervals  $[x_{i-1}, x_i]$  (i.e.  $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ ), while  $f_i \equiv f(\xi_i)$  with  $\xi_i$  some point in sub-interval  $i$ .

A sufficient condition for the *Riemann integral* to exist for the function  $f$  is that any such sum (any choice of  $x_i$  [for which  $\max_i |x_i - x_{i-1}| \rightarrow 0$ ] and  $\xi_i$ ) converge to the same value in the limit  $n \rightarrow \infty$  [21]. In this case we say

$$\lim_{n \rightarrow \infty} S_n = S = \int_a^b f(x) dx, \quad (\text{B.2})$$

and that  $f$  is Riemann integrable.

A less strict, but still sufficient condition is to chose  $\overline{f_i} = \max\{f(x) : x \in [x_{i-1}, x_i]\}$  and  $\underline{f_i} = \min\{f(x) : x \in [x_{i-1}, x_i]\}$  and then only demand that the two sums converge to a common limit [22],

$$\begin{aligned} \lim_{n \rightarrow \infty} \overline{S_n} &= \lim_{n \rightarrow \infty} \sum_{i=1}^n (x_i - x_{i-1}) \overline{f_i} \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n (x_i - x_{i-1}) \underline{f_i} = \lim_{n \rightarrow \infty} S_n \\ &= \int_a^b f(x) dx. \end{aligned}$$

Although easier than checking *every possible* Riemann sum, checking that the two upper and lower sums converge to a common limit is still a somewhat tedious

procedure for checking integrability. In fact it turns out that a sufficient condition on  $f$  is that it is continuous and bounded on  $[a, b]$  [21]. The latter condition is not necessary on a finite interval since all continuous functions on a closed finite domain are bounded according to the extreme value theorem. However, if we extend the limits of integration to an infinite interval, for example  $[0, \infty)$ , then the boundedness of  $f$  is not guaranteed by the continuity we need to explicitly demand  $|f(x)| < \infty$  for all  $x \in [a, b]$ .

It is easy to see that the converse is *not* true. Any Riemann integrable function is not automatically continuous. Take for example the integral over  $[0, 1]$  with the step function

$$f(x) = \begin{cases} 1 & \text{if } x > 1/2 \\ 0 & \text{else} \end{cases}. \quad (\text{B.3})$$

Even though the upper and lower Riemann sums both attain the value  $1/2$  in the limit  $n \rightarrow \infty$  and the function is Riemann integrable, it demonstrably is not continuous. A more careful analysis shows that a less strict but sufficient condition on  $f$  is that it be continuous *almost everywhere* on  $[a, b]$  (i.e. continuous on all of the interval, except possibly on a subset  $C \subset [a, b]$  with measure zero) [23]. With this condition, the converse also holds.

### Newton-Cotes quadrature

Since the Riemann integral is defined in terms of the limit of a sum, numerical approximations to it arise naturally from any scheme for choosing  $\xi_i$  and the partitioning. One of the simplest possible approximations is to take the midpoint value of each sub-interval to be  $\xi_i$  with a uniform mesh of equispaced  $x_i$ s. This constitutes the **midpoint rule** [21],

$$I \approx \sum_{i=1}^n f(x_{i-1} + \Delta x/2)(x_i - x_{i-1}) = \Delta x \sum_{i=1}^n f(x_{i-1} + \Delta x/2), \quad (\text{B.4})$$

where  $\Delta x \equiv (x_i - x_{i-1})$  which is the same for all  $i$ .

Instead of the midpoint, we can use the *average* of the left and right endpoints of the subinterval as  $\xi_i$ . Geometrically, this means we are approximating the integral of each sub-interval by the integral over a right trapezoid with base points at  $(x_{i-1}, 0)$  and  $(x_i, 0)$  and upper point at the function values  $(x_{i-1}, f(x_{i-1}))$  and  $(x_i, f(x_i))$ . The resulting approximation is known as the **trapezoidal rule** [11],

$$I \approx \sum_{i=1}^n \frac{f(x_{i-1}) + f(x_i)}{2}(x_i - x_{i-1}) = \Delta x \sum_{i=1}^n \frac{f(x_{i-1}) + f(x_i)}{2}. \quad (\text{B.5})$$

Yet another numerical scheme arises from replacing the integrand in each sub-interval with an interpolating polynomial of degree two, which by construction coincides with  $f$  at the endpoints and the midpoint. This constitutes **Simpson's rule**

[21],

$$\begin{aligned}
 I &\approx \sum_{i=1}^n (x_i - x_{i-1}) \left( \frac{f(x_{i-1})}{6} + \frac{4f(x_{i-1} + \Delta x/2)}{6} + \frac{f(x_i)}{6} \right) \\
 &= \frac{\Delta x}{6} \sum_{i=1}^n \left( f(x_{i-1}) + 4f(x_{i-1} + \Delta x/2) + f(x_i) \right). \tag{B.6}
 \end{aligned}$$

All three approximations are examples of Newton-Cotes quadrature rules, which approximate the integral by replacing the integrand by interpolating polynomials of order  $k$  on each of the  $n$  sub-intervals. We can build arbitrarily high order methods by constructing higher order interpolating polynomials within each interval. The interpolation procedure is described for example in [24]. We have just seen Newton-Cotes method for orders zero (midpoint rule, zero order polynomial [constant]), one (trapezoidal rule, linear polynomial), and two (Simpson's rule, quadratic polynomial). The next few commonly used methods are the third order *Simpson's 3/8 rule* and the fourth order *Boole's rule*.

### Gaussian quadrature

Note that so far we have assumed the sub-intervals to all be the same size. If we drop this requirement, we can construct more advanced rules which exploit some convenient properties of orthogonal polynomials. Gaussian quadrature rules are a set of schema for numerical intergration in which we extract a *weight function* from the integrand

$$\int_a^b f(x) dx = \int_a^b W(x)g(x) dx \approx \sum_{i=1}^n w_i g(x_i). \tag{B.7}$$

The weight function is associated with a set of orthogonal polynomials, and the integration points  $x_i$  are chosen as the zeros of the polynomial of degree  $n - 1$ . Note carefully that  $w_i \neq W(x_i)$ . The weights  $w_i$  can in general be expressed as [25]

$$w_i = \left( \frac{a_n}{a_{n-1}} \right) \frac{\int_a^b W(x)p_{n-1}(x)^2 dx}{p'_n(x_i)p_{n-1}(x_i)} \tag{B.8}$$

where  $p_n(x)$  is the orthogonal polynomial of degree  $n$  and  $a_n$  is the coefficient of the  $x^n$  term in  $p_n(x)$ . In some cases, the weight function is present in the original integral and the extraction constitutes a strict simplification of the function. For example, with Chebyshev polynomials<sup>1</sup> the weight function takes the form  $W(x) =$

---

<sup>1</sup>The Chebyshev polynomials are solutions to the differential equation

$$(1 - x^2) \frac{\partial^2 y(x)}{\partial x^2} - x \frac{\partial y(x)}{\partial x} + n^2 y(x) = 0, \tag{B.9}$$

$1/\sqrt{1-x^2}$ , so trying to apply Gauss-Chebyshev quadrature to the integrand  $(x^{10} + x + 2)/\sqrt{1-x^2}$  would yield simply  $g(x) = x^{10} + x + 2$  and we would just have to evaluate  $g_i$  according to the zeroes of the  $n$ th Chebyshev polynomial. Each class of polynomials is associated with a specific interval of integration. For Chebyshev, this is  $[-1, 1]$ . So using our previous example, we note that with only  $3!$  integration points (exclamation point for emphasis *and* factorial function) we integrate *exactly*

$$I \equiv \int_{-1}^1 \underbrace{\frac{x^{10} + x + 2}{\sqrt{1-x^2}}}_{\equiv f(x)} dx = \sum_{i=1}^6 w_i \underbrace{(x^{10} + x + 2)}_{g(x)} = \frac{575\pi}{256}. \quad (\text{B.11})$$

In general, if  $g(x)$  is a polynomial of degree  $2n - 1$  for a weight function associated with some class of orthogonal polynomials, then the gaussian quadrature rule associated with the same class of polynomials will integrate the original  $f(x)$  (recall that  $g(x) = f(x)/W(x)$ ) *exactly* with only  $n$  integration points [11].

### Multiple integrals

Both of the aforementioned rules are straight forward to extend to higher dimensional integrals. For the Newton-Cotes rules, we can simply apply the rule again to the sum resulting from the application of the rule, i.e.

$$\begin{aligned} I_{2D} &= \int_a^b \int_a^b f(x, y) dx dy \approx \int_a^b \sum_{i=1}^n \Delta x f(\xi_i, y) dy \\ &\approx \sum_{i=1}^n \sum_{j=1}^n \Delta x \Delta y f(\xi_i, \zeta_j). \end{aligned} \quad (\text{B.12})$$

Since function evaluations on the endpoints of sub-intervals (sub-areas to be precise) coincide with the endpoints of the neighbouring sub-intervals, a number of points may be evaluated multiple times and thus have a higher *weight* in the final sum. For example, the 1D trapezoidal rule carries weights

$$1/2 \quad 1 \quad 1 \quad 1 \quad \dots \quad 1 \quad 1 \quad 1 \quad 1/2, \quad (\text{B.13})$$

---

with  $n$  a non-negative integer. In general, the solution can be written as [26]

$$T_n(x) = \sum_{k=0}^{\lfloor 1/2 \rfloor} \binom{n}{2k} (x^2 - 1) x^{n-2k}. \quad (\text{B.10})$$



since

$$\begin{aligned}
\frac{\Delta x}{2} \sum_{i=1}^n \left( f(x_{i-1}) + f(x_i) \right) &= \frac{\Delta x}{2} \left[ f(x_0) + 2 \sum_{i=1}^{n-1} \left( f(x_i) + f(x_{i+1}) \right) \right] \\
&= \frac{\Delta x}{2} \left[ f(x_0) + 2 \left( \sum_{i=1}^{n-1} f(x_i) \right) + f(x_n) \right] \\
&= \frac{\Delta x}{2} [f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-2}) + 2f(x_{n-1}) + f(x_n)].
\end{aligned} \tag{B.14}$$

In a similar way, the 2D trapezoidal rule has the weights

$$\begin{array}{cccccccccc}
1/2 & 1 & 1 & 1 & \dots & 1 & 1 & 1 & 1/2 & \\
1 & 2 & 2 & 2 & \dots & 2 & 2 & 2 & 1 & \\
1 & 2 & 2 & 2 & \dots & 2 & 2 & 2 & 1 & \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \\
1 & 2 & 2 & 2 & \dots & 2 & 2 & 2 & 1 & \\
1 & 2 & 2 & 2 & \dots & 2 & 2 & 2 & 1 & \\
1/2 & 1 & 1 & 1 & \dots & 1 & 1 & 1 & 1/2 & ,
\end{array} \tag{B.15}$$

while the 2D Simpson's rule attains the weights (apart from a factor  $1/6$ )

$$\begin{array}{cccccccccccc}
1 & 4 & 2 & 4 & 2 & \dots & 2 & 4 & 2 & 4 & 1 & \\
4 & 16 & 8 & 16 & 8 & \dots & 8 & 16 & 8 & 16 & 4 & \\
2 & 8 & 4 & 8 & 4 & \dots & 4 & 8 & 4 & 8 & 2 & \\
4 & 16 & 8 & 16 & 8 & \dots & 8 & 16 & 8 & 16 & 4 & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
4 & 16 & 8 & 16 & 8 & \dots & 8 & 16 & 8 & 16 & 4 & \\
2 & 8 & 4 & 8 & 4 & \dots & 4 & 8 & 4 & 8 & 2 & \\
4 & 16 & 8 & 16 & 8 & \dots & 8 & 16 & 8 & 16 & 4 & \\
1 & 1 & 2 & 4 & 2 & \dots & 2 & 4 & 2 & 4 & 1 & .
\end{array} \tag{B.16}$$

A similar scheme yields multi-dimensional Gaussian quadrature rules, where the total weights become products of the 1D weights.



# Appendix C

## Functionals and functional variations

Recall that a *function* is a mapping from some algebraic scalar field  $\mathbb{F}$  to another (possibly different) field  $\mathbb{F}'$ , i.e.  $g : \mathbb{F} \rightarrow \mathbb{F}'$ . In physics, we are usually interested mainly in the cases where  $\mathbb{F}$  and  $\mathbb{F}'$  are the real or complex numbers,  $\mathbb{R}$  or  $\mathbb{C}$ . An example is the complex exponential,  $x \mapsto e^{ix}$ , which takes complex values but the argument is real, so  $g : \mathbb{R} \rightarrow \mathbb{C}$  in this case.

A *functional*, on the other hand, is a mapping from some function space,  $\mathcal{F}$ , to a scalar field  $\mathbb{F}$ , i.e.  $f : \mathcal{F} \rightarrow \mathbb{F}$ . We will take the space of functions to be the underlying Hilbert space of our quantum mechanical system,  $\mathcal{H}$ , and the field to be the complex numbers,  $\mathbb{C}$ . The functional thus assigns to each  $f \in \mathcal{H}$  a complex number.

As a familiar example of such a construction, let us consider the definite integral. For the moment, let us take the function space to be continuous real functions of a single real argument in the range  $[0, 1]$ ,  $C([0, 1])$ . We call this *functional*  $I$ , such that  $I : C([0, 1]) \rightarrow \mathbb{R}$ .

$$I[f] = \int_0^1 dx f(x) \quad (\text{C.1})$$

thus assigns a real number to any continuous function on  $[0, 1]$ . For example,  $I[e^x] = e - 1 \approx 1.7183$  or  $I[\sqrt{x}] = 2/3 \approx 0.6667$ .

When working in a separable Hilbert space as we always do in quantum mechanics, we may always express any function  $f \in \mathcal{H}$  in terms of some basis  $\{ \chi_n \}_{n=1}^\infty$ , (recall the Parseval relation from section ((QM math)))

$$|f\rangle = \left( \sum_{n=1}^\infty |\chi_n\rangle \langle \chi_n| \right) |f\rangle = \sum_{n=1}^\infty \underbrace{\langle \chi_n | f \rangle}_{c_n} |\chi_n\rangle = \sum_{n=1}^\infty c_n |\chi_n\rangle, \quad (\text{C.2})$$

meaning we can think of a functional  $F[f]$  as a *function* of the vector of coefficients relative to the basis set,  $\mathbf{c} = (c_1, c_2, \dots)$  [9].

### C.0.1 Short mathematical interlude

Let  $B(X, Y)$  denote the set of all continuous linear transformations from normed vector spaces  $X$  and  $Y$  (over the algebraic scalar field  $\mathbb{F}^1$ ). For example we may consider  $X = \mathbb{R}^n$  and  $Y = \mathbb{R}^m$ , i.e. the set of real vectors of length  $n$  and  $m$ , respectively. The set of continuous linear transformations from  $X$  to  $Y$ ,  $B(X, Y)$ , thus consists of real valued matrices of dimensions  $m \times n$ , so we may write  $B(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{m \times n}$ .

A *Banach space* is a normed vector space which is complete under the metric associated with the norm. Since any norm,  $\|\cdot\|$  induces a metric by  $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ , and the inner product  $\langle \cdot | \cdot \rangle$  induces a norm by  $\|\cdot\| = \sqrt{\langle \cdot | \cdot \rangle}$  we can define a *Hilbert space* as a Banach space which is complete w.r.t. this specific metric [22, 23].

The space of linear transformations from  $X$  to  $\mathbb{F}$ , with  $X$  being some normed vector space, is called the *dual space* of  $X$ , sometimes denoted  $X^*$ . We note that a linear transformation from  $X$  to  $\mathbb{F}$  is exactly a linear *functional*, and so functionals "live in the dual" of the vector space itself. It turns out that if  $X$  is normed, the dual is always a Banach space [27]. Since we are inherently working with Hilbert spaces in quantum mechanics, it is natural to ask: What can we say in general about the vector space of functionals on a Hilbert space  $\mathcal{H}$ ?

In the following, we take  $f \in \mathcal{H}^*$  to be a linear functional on  $\mathcal{H}$  and  $x \in \mathcal{H}$  to be a function in the Hilbert space itself. It can be shown that for any such  $x$  there exists a *unique*  $y \in \mathcal{H}$  such that  $f[x] = f_y[x] = \langle x | y \rangle$ , where the functional  $f_y[\cdot] \equiv \langle \cdot | y \rangle$  [27, 23]. This is known as the Riesz representation theorem or sometimes the Riesz-Fréchet theorem. Essentially, this means that we can associate the dual space of  $\mathcal{H}$  with the space itself since there is a correspondance between the functionals and the elements of the space itself. This is more succinctly stated as Hilbert spaces are *self-dual*, and it is this property that justifies the use of Dirac bra-ket notation since we are guaranteed that any ket has a unique corresponding bra which is its Hermitian conjugate.

### C.0.2 Functional differentials and derivatives

The differential of a functional  $F[f]$  is the part of the difference  $F[f + \delta f] - F[f]$  that depends linearly on  $\delta f$ , where  $\delta f$  is an infinitesimal variation of the argument function  $f$  [28]. Since we need to account for the continuous variation of  $F$  over the infinitesimal range  $[f, f + \delta f]$  we take the integral

$$\delta F[f] = \int \frac{\delta F[f]}{\delta f(x)} \delta f(x) dx, \quad (\text{C.3})$$

where we have defined the *functional derivative* of  $F$  w.r.t.  $f$  at the point  $x$  as

$$F'[f] \equiv \frac{\delta F[f]}{\delta f(x)}. \quad (\text{C.4})$$

---

<sup>1</sup>Meaning  $X$  and  $Y$  are closed under scalar multiplication with elements  $c \in \mathbb{F}$ .

If the underlying space is a Banach space, meaning the dual space is also a Banach space (c.f. section C.0.1), we can write the functional differential in a way that is familiar [29]:

$$\delta F[f] = \lim_{\varepsilon \rightarrow 0} \frac{F[f + \varepsilon \delta f(x)] - F[f]}{\varepsilon} = \int \frac{\delta F[f]}{\delta f(x)} \delta f(x) dx. \quad (\text{C.5})$$

In the following, assume  $g[f]$  is a functional of the function  $f$ . It turns out that the functional derivative behaves a lot like ordinary derivatives, [30]

$$\frac{\delta}{\delta f(x)} (aF[f] + bG[f]) = a \frac{\delta F[f]}{\delta f(x)} + b \frac{\delta G[f]}{\delta f(x)} \quad (\text{linearity})$$

$$\frac{\delta}{\delta f(x)} (F[f] G[f]) = G[f] \frac{\delta F[f]}{\delta f(x)} + F[f] \frac{\delta G[f]}{\delta f(x)} \quad (\text{product rule})$$

$$\frac{\delta}{\delta f(x)} (F[g]) = \int \frac{\delta F[g]}{\delta g(x')} \frac{\delta g(x')}{\delta f(x)} dx \quad (\text{chain rule})$$

We can also define higher-order functional derivatives, for example the equivalent of the ordinary double derivative

$$\frac{\delta^2 F[f]}{\delta f(x) \delta f(x')} = \frac{\delta}{\delta f(x)} \left( \frac{\delta F[f]}{\delta f(x')} \right). \quad (\text{C.6})$$

We may use this to compute the Taylor expansion of a functional  $F[f]$  as

$$\begin{aligned} F[f + \Delta f] &= F[f] + \sum_{n=1}^{\infty} \frac{1}{n!} \int \cdots \int \frac{\delta^{(n)} F[f]}{\delta f(x_1) \cdots \delta f(x_n)} \Delta f(x_1) \Delta f(x_2) \cdots \Delta f(x_n) dx_1 dx_2 \cdots dx_n \\ F[f + \Delta f] &= F[f] + \int \frac{\delta F[f]}{\delta f(x)} \Delta f(x) dx + \frac{1}{2} \int \int \frac{\delta^2 F[f]}{\delta f(x) \delta f(x')} \Delta f(x) \Delta f(x') dx dx' + \cdots, \end{aligned} \quad (\text{C.7})$$

where  $\Delta f(x)$  is a finite (not infinitesimal) variation in the function  $f(x)$  [28].



# Bibliography

- [1] Wei Hu, Lin Lin, and Chao Yang. “DGDFT: A massively parallel method for large scale density functional theory calculations”. In: *The Journal of Chemical Physics* 143.12 (2015), p. 124110. DOI: [10.1063/1.4931732](https://doi.org/10.1063/1.4931732). eprint: <https://doi.org/10.1063/1.4931732>. URL: <https://doi.org/10.1063/1.4931732>.
- [2] D.R. Bowler and T. Miyazaki. “Calculations for millions of atoms with density functional theory: linear scaling shows its potential”. In: *Journal of Physics: Condensed Matter* 22.7 (2010), p. 074207. URL: <http://stacks.iop.org/0953-8984/22/i=7/a=074207>.
- [3] Joost VandeVondele, Urban Borštnik, and Jürg Hutter. “Linear Scaling Self-Consistent Field Calculations with Millions of Atoms in the Condensed Phase”. In: *Journal of Chemical Theory and Computation* 8.10 (2012). PMID: 26593003, pp. 3565–3573. DOI: [10.1021/ct200897x](https://doi.org/10.1021/ct200897x). eprint: <http://dx.doi.org/10.1021/ct200897x>. URL: <http://dx.doi.org/10.1021/ct200897x>.
- [4] Gongpu Zhao et al. “Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics”. In: *Nature* 497 (May 2013), URL: <http://dx.doi.org/10.1038/nature12162>.
- [5] Tyler Reddy et al. “Nothing to Sneeze At: A Dynamic and Integrative Computational Model of an Influenza A Virion”. In: *Structure* 23.3 (2015), pp. 584–597. ISSN: 0969-2126. DOI: <https://doi.org/10.1016/j.str.2014.12.019>. URL: <https://www.sciencedirect.com/science/article/pii/S0969212615000325>.
- [6] R. E. Angulo et al. “Scaling relations for galaxy clusters in the Millennium-XXL simulation”. In: *Monthly Notices of the Royal Astronomical Society* 426.3 (2012), pp. 2046–2062. DOI: [10.1111/j.1365-2966.2012.21830.x](https://doi.org/10.1111/j.1365-2966.2012.21830.x). eprint: [/oup/backfile/content\\_public/journal/mnras/426/3/10.1111/j.1365-2966.2012.21830.x/2/426-3-2046.pdf](http://oup/backfile/content_public/journal/mnras/426/3/10.1111/j.1365-2966.2012.21830.x/2/426-3-2046.pdf). URL: <http://dx.doi.org/10.1111/j.1365-2966.2012.21830.x>.
- [7] Juhan Kim et al. “The new horizon run cosmological N-body simulations”. In: *arXiv preprint arXiv:1112.1754* (2011).
- [8] T. Helgaker, P. Jørgensen, and J. Olsen. *Molecular Electronic-Structure Theory*. John Wiley & Sons, 2000. ISBN: 0-471-96755-6.
- [9] S. Kvaal. *Lecture notes for FYS-KjM4480*. Lecture notes. Sept. 2017. URL: <http://theory.rutgers.edu/~giese/notes/DFT.pdf>.

- [10] I. Shavitt and R.J. Bartlett. *Many-Body Methods in Chemistry and Physics. MBPT and Coupled-Cluster Theory*. Cambridge University Press, 2009. ISBN: 0-521-81832-X.
- [11] M. Hjorth-Jensen. *Computational Physics*. Lecture notes. Aug. 2015. URL: <https://github.com/CompPhysics/ComputationalPhysics/blob/master/doc/Lectures/lectures2015.pdf>.
- [12] B.L. Hammond, W.A. Lester Jr., and P.J. Reynolds. *Monte Carlo Methods in Ab Initio Quantum Chemistry*. Wspc, 1994. ISBN: 981-02-0322-5.
- [13] Harold Conroy. “Molecular Schrödinger Equation. II. Monte Carlo Evaluation of Integrals”. In: *The Journal of Chemical Physics* 41.5 (1964), pp. 1331–1335. DOI: [10.1063/1.1726069](https://doi.org/10.1063/1.1726069). eprint: <https://doi.org/10.1063/1.1726069>. URL: <https://doi.org/10.1063/1.1726069>.
- [14] J.B. Anderson. *Quantum Monte Carlo. Origins, Development, Applications*. Oxford University Press, 2007. ISBN: 0-19-531010-1.
- [15] D. R. Hartree. “The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 24.1 (1928), pp. 89–110. DOI: [10.1017/S0305004100011919](https://doi.org/10.1017/S0305004100011919).
- [16] V. Fock. “Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems”. In: *Zeitschrift für Physik* 61.1 (Jan. 1930), pp. 126–148. ISSN: 0044-3328. DOI: [10.1007/BF01340294](https://doi.org/10.1007/BF01340294). URL: <https://doi.org/10.1007/BF01340294>.
- [17] A. Szabo and N.S. Ostlund. *Modern Quantum Chemistry*. Dover Publications, 1996. ISBN: 0-486-69186-1.
- [18] W. Kohn and L. J. Sham. “Self-Consistent Equations Including Exchange and Correlation Effects”. In: *Phys. Rev.* 140 (4A Nov. 1965), A1133–A1138. DOI: [10.1103/PhysRev.140.A1133](https://doi.org/10.1103/PhysRev.140.A1133). URL: <https://link.aps.org/doi/10.1103/PhysRev.140.A1133>.
- [19] R.M. Martin. *Electronic Structure*. Cambridge University Press, 2004. ISBN: 0-521-53440-2.
- [20] Axel D. Becke. “Perspective: Fifty years of density-functional theory in chemical physics”. In: *The Journal of Chemical Physics* 140.18 (2014), 18A301. DOI: [10.1063/1.4869598](https://doi.org/10.1063/1.4869598). eprint: <https://doi.org/10.1063/1.4869598>. URL: <https://doi.org/10.1063/1.4869598>.
- [21] P.J. Davis and P. Rabinowitz. *Methods of Numerical Integration*. 2nd ed. Academic Press, 1984. ISBN: 0-12-206360-0.
- [22] T. Lindstrøm. *Mathematical Analysis*. Lecture notes. 2017. URL: <http://folk.uio.no/snorrec/17VMat2400/Spaces.pdf>.
- [23] J.M. McDonald and N.A. Weiss. *A course in Real Analysis*. 2nd ed. Academic Press, 2013. ISBN: 0-12-387774-1.



- [24] K. Mørken. *Numerical Algorithms and Digital Representation*. Lecture notes. Sept. 2017. URL: <http://www.uio.no/studier/emner/matnat/math/MAT-INF1100/h17/kompendiet/matinf1100.pdf>.
- [25] V.I. Krylov. *Approximate Calculation of Integrals*. 1st ed. Dover Publications, 2005. ISBN: 0-486-44579-8.
- [26] K. Rottmann. *Matematisk Formelsamling*. 10th ed. Spektrum Forlag, 2008. ISBN: 8278220050.
- [27] B.P. Rynne and M.A. Youngson. *Linear Functional Analysis*. 2nd ed. Springer-Verlag London, 2008. ISBN: 1-84800-004-9.
- [28] R.G. Parr and W. Yang. *Density-functional theory of atoms and molecules*. Oxford University Press, 1989. ISBN: 0-19-504279-4.
- [29] P. Echenique and J.L. Alonso. “A mathematical and computational review of Hartree-Fock SCF methods in quantum chemistry”. In: *Molecular Physics* 105.23-24 (2007), pp. 3057–3098. DOI: [10.1080/00268970701757875](https://doi.org/10.1080/00268970701757875). eprint: <http://dx.doi.org/10.1080/00268970701757875>. URL: <http://dx.doi.org/10.1080/00268970701757875>.
- [30] J. Toulouse. *Introduction to density-functional theory*. Lecture notes. June 2017. URL: [http://www.lct.jussieu.fr/pagesperso/toulouse/enseignement/introduction\\_dft.pdf](http://www.lct.jussieu.fr/pagesperso/toulouse/enseignement/introduction_dft.pdf).
- [31] *Quantities and units – Part 2: Mathematical signs and symbols to be used in the natural sciences and technology*. Standard. Geneva, CH: International Organization for Standardization, Dec. 2009.