1
Introduction

> *The book of nature you can only understand, if you have previously learnt its language and the letters. It is written in mathematical language and the letters are geometrical figures, and without these means it is impossible for human beings to understand even a word of it.*
>
> Galileo Galilei, 16th century

1.1
Theory and Models – Interpretation of Experimental Data

Every experimental result is – at the best – as good as the theoretical model used for its interpretation.... This sentence has been chosen as a leading remark for this book, as chemistry is still largely a field of science dominated by more or less simple models employed to 'explain' (rather: describe) the behaviour of molecules, their interactions and reactions, and to interpret all the data acquired by sophisticated measurements. The terminology introduced by these models is very rarely questioned for its validity and compatibility with state-of-the-art theoretical knowledge, and thus many of the interpretations of experimental data and subsequent conclusions could be inappropriate or even erroneous. On the other hand, a valid theoretical background for chemistry has been existent since 1926, when Erwin Schrödinger formulated his famous equation. The reason why theory has not yet penetrated chemistry as it did physics during the early 20th century is to be seen in the inability to solve Schrödinger's equation by analytical methods for other than one-electron systems. To date, only numerical solutions are possible, and one had to wait for the capacity of high-performance computers to deal with chemically relevant systems to test, prove, and successfully apply theoretical chemistry methods. This capability has been reached only recently – for a larger scientific community during the past two decades – and it can be expected, therefore, that the 21st century will be the one in which theory is given a much greater role in chemistry, reforming curricula and research practice similar to the development of physics some 100 years ago.

The Basics of Theoretical and Computational Chemistry. Edited by B. M. Rode, T. S. Hofer and M. D. Kugler

ISBN: 978-3-527-31773-8

Today, computational chemistry methods are almost ubiquitously used, and publications frequently contain 'theoretical' sections which are often nothing more than black-box applications of commercial program packages. Such an inappropriate use of theoretical methods is often caused by a lack of knowledge of the quantum theoretical foundations of the available programs and a simultaneous adherence to simplistic models employed to interpret the quantitative results obtained by the computations in qualitative terms of these models.

The purpose of this book is to familiarise not only chemistry students, but also those chemists who did not have the chance to obtain a good theoretical background during their studies, with the basics of theoretical and computational chemistry. To an extent, this should enable them to understand the underlying physical principles, to judge the validity of the commonly used models, and to obtain sufficient knowledge to use computational chemistry methods in a prudent and appropriate way. In order to achieve these goals, mathematical requirements have been reduced to a minimum, without sacrificing physical rigour in the theoretical framework. This could be realised by using the vector space theory of matter with its largely linear algebraic formalism instead of the commonly employed formalism of integro-differential equations following the historical development of quantum chemistry. Detailed descriptions of procedures that might be important for the specialist – but not for the generally interested reader – have been omitted for the sake of clarity and conciseness. In the final sections, the basic principles of perturbation theory and group theory are outlined with regard to their use in chemistry, and a brief overview of the most important methods in contemporary computational chemistry is provided. Thus, the book should serve as a good general introduction into theory of chemistry and create a good basis for further, more specialised reading, wherever this is needed or desired.

Before starting the first chapters of this book, it seems important to attempt a definition of the difference between 'models' and 'theory'. A 'model' can be considered a fiction with a certain power to indicate some aspects of 'reality' and to rationalise a (limited) number of observations, sometimes also allowing some predictions. A 'theory' should allow a general and precise explanation/description of all (or at least as many as possible) phenomena and a reliable prediction of the results of any future observation. In this case, the fewer postulates needed in its formalism and the wider its field of validity, the better the theory. Chemistry is in the somewhat fortunate situation of being mainly concerned with very few of the manifold elementary particles, namely with electrons, nuclei (which can mostly be treated as charged masses of virtually no dimension), and photons. During the course of this book it will be seen that a theoretical concept for this case can be built on three well-proven postulates and two empirical observations, and that all further aspects of the theory result as logical consequences, thus making the development of the theory a straightforward procedure. Figure 1.1 provides a schematic overview of the general concept followed in this book.

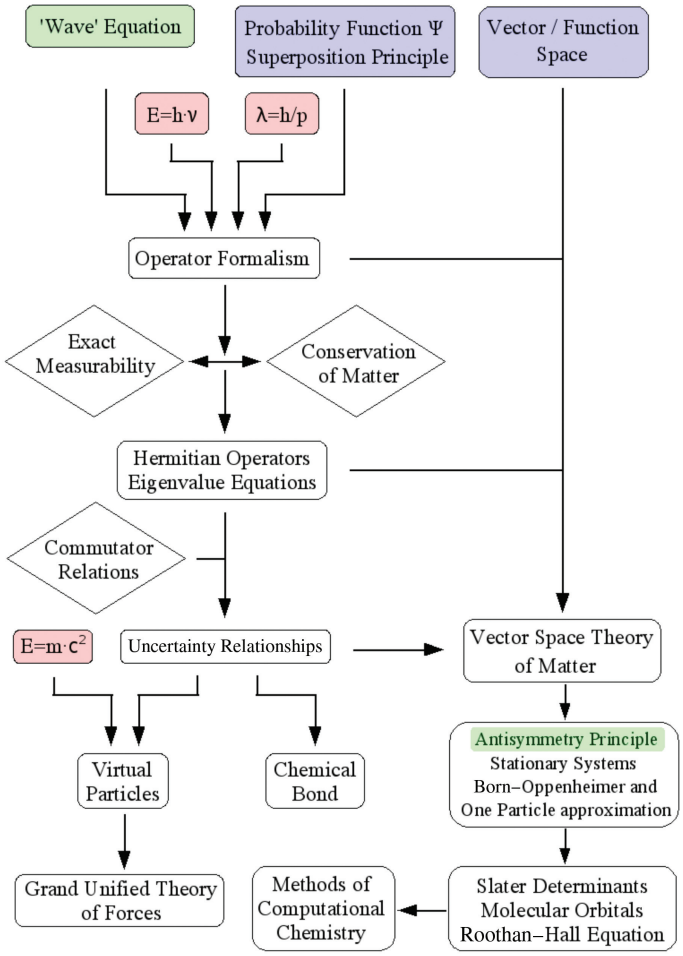

Fig. 1.1 'Float Chart' of the concept followed by this book. Postulates are indicated by red, empirical observations by green, and mathematical concepts by blue.

1.2 The Notation

To facilitate reading of this book, the consistent notation used in it will be summarised here, defining symbols and associated meanings. The following list should also serve as a convenient tool to identify any specific notation used in the text.

$|x\rangle$ denotes a vector in the 'ket' space, $\langle x|$ denotes the complex conjugated and transposed of $|x\rangle$, and thus a vector of the 'bra' space.

The associated sets of vectors are called 'bra' and 'ket', because $\langle\ \rangle$ corresponds to a 'bracket'. This notation is also known as *Dirac* notation. In detail, the vectors are defined as follows:

$$|x\rangle = (x_1, x_2, \ldots) \qquad \langle x| = \begin{pmatrix} x_1^* \\ x_2^* \\ \cdots \end{pmatrix}$$

φ denotes a function, and is usually also written as a vector $|\varphi\rangle$, due to the essential equivalence of a function with a vector: if we write the values of a function at subsequent values of its variable, we obtain an n-tuple of numbers, i.e., a vector ('digitised function').

$\mathbb{A}$ or $\mathbb{S}$ ('fat symbols') are used for matrices, and this is equally applied to Greek letters: $\mathbf{\Phi}$ and $\mathbf{\Psi}$ denotes matrices.

The determinant of a matrix is given by $|\mathbb{A}| = \det(\mathbb{A})$. The transposed matrix $\mathbb{A}$ is written as $\mathbb{A}^T$, the adjugate matrix as $\mathbb{A}^+$ and the inverse as $\mathbb{A}^{-1}$.

$$\mathbb{1} = \begin{pmatrix} 1 & \cdots & \emptyset \\ \emptyset & \cdots & \emptyset \\ \cdots & \cdots & \cdots \\ \emptyset & \cdots & 1 \end{pmatrix}$$

is the unit matrix and $\mathbb{0}$ the zero-matrix (containing only $\emptyset$ values).

Vectors and functions are often constructed from sets of more or less simple 'basis' vectors/functions. $\{e_i\}^n$ and $\{\varphi_i\}^n$ indicate such basis sets.

Operators will be marked with a hat; for example, $\hat{H}$ stands for the Hamilton operator. Special operators are the direct-sum $\oplus$ and the direct-product $\otimes$ operators and the $\hat{\mathbf{1}}$ (identity) and $\hat{\mathbf{0}}$ (annihilation) operators.

Two other frequently used special operators are the Nabla and the Laplace operator:

$$\text{Nabla operator:} \quad \nabla = \frac{\partial}{\partial x}|i\rangle + \frac{\partial}{\partial y}|j\rangle + \frac{\partial}{\partial z}|k\rangle$$

$$\text{Laplace operator:} \quad \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

The Nabla operator is a vector operator, while the Laplace operator is scalar, resulting from the scalar product of the Nabla Operator with itself: $\Delta = \langle\nabla|\nabla\rangle$.

The Kronecker–Delta is another important notation, meaning:

$$\delta_{ij} = \begin{cases} 1 & \forall i = j \\ \emptyset & \text{otherwise} \end{cases}$$

The following symbols are also used in this book with the meanings:

$\exists$ ... there is

$\forall$ ... for all

$\in$ ... is an element of

iff ... if and only if

$d\tau = dx\,dy\,dz$

1.3 Vector Space V^n and Function Space F^n

V^n denotes a vector space of the dimension n. Thus, V^2 is a two-dimensional vector space (i.e., a plane) and V^3 is a three-dimensional vector space, corresponding to our familiar x/y/z coordinate system. To describe a vector space, one needs 'basis vectors': n linearly independent vectors span an n-dimensional vector space. Linear independence means that none of these basis vectors can be constructed by a linear combination of the other $n - 1$ basis vectors. If all basis vectors are perpendicular to each other, they form an orthogonal basis set.

Basis sets are denoted as $\{e_i\}^n$ for V^n. In Cartesian coordinates the basis for the V^3 consists of three basis vectors in x-, y-, and z-direction:

$$\{e_i\}^3 = (|i\rangle, |j\rangle, |k\rangle)$$

The function space F^n is the analogue of the vector space, using n linearly independent functions φ_i as basis $\{\varphi_i\}^n$.

Basically, a function is a transformation from the variable x to its function value $f(x)$. In fact it is a continuous transformation, but if one digitizes the function in steps Δx, a finite vector is obtained instead of the function. So, x is reproduced as f_1, $x + \Delta x$ as f_2 and so on, resulting in an n-tuple of numbers $(f_1, f_2, \ldots)$. If Δx decreases toward zero, the dimension of the vectors grows, finally reaching infinity, i.e., a vector in V^∞.

The equivalence between vector and function space will be used throughout this book, as it allows a much more convenient and easier-to-manipulate formulation of equations, which can easily be transformed into computer language.

The Scalar Product

Given, the vectors $|a\rangle$ and $|b\rangle$, the scalar product $\langle a|b\rangle$ is obtained by $\sum_i a_i^* \cdot b_i$. The analogue for functions is a sum with infinitesimal step width, i.e., an integral. Thus, $\langle\varphi_1|\varphi_2\rangle$ corresponds to $\int \varphi_1^* \cdot \varphi_2\, d\tau$. We will make continuous use of

this simplified notation of integrals in function space as scalar products in vector space.

1.4 Linear Transformation – Change of Basis

Often it is an advantage to change bases. To change from a basis $\{e_i\}^n = (e_1 \dots e_n)$ to another basis of the same dimension n, $\{x_i\}^n = (x_1 \dots x_n)$, one performs a linear tranformation. As an example we will consider a vector $|u\rangle$, represented in both bases.

Let $|u\rangle$ be an element of V^2, given in $\{e_i\}^2$ as $|u\rangle^e = 3|e_1\rangle + 2|e_2\rangle$. We wish to obtain its representation in $\{x_i\}^2$, namely $|u\rangle^x$. For this purpose, we must know the relationship between the two bases. As every vector of V^2 can be constructed from a given basis set, the two vectors forming the basis $\{x_i\}^2$ can also be constructed from the basis vectors of $\{e_i\}^2$. In our example, let the connection between the different basis vectors be given by:

$$|x_1\rangle = 2|e_1\rangle - |e_2\rangle \qquad \text{and} \qquad |x_2\rangle = |e_1\rangle + 3|e_2\rangle.$$

In matrix notation this reads $\mathbb{X} = \mathbb{E} \cdot \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix} = \mathbb{E} \cdot \mathbb{T}$, or $\mathbb{X}\mathbb{T}^{-1} = \mathbb{E}$, where $\mathbb{T}$ is the transformation matrix from $\mathbb{E}$ to $\mathbb{X}$.

We can now write

$$\mathbb{U} = \mathbb{E}\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \mathbb{X}\mathbb{T}^{-1}\begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

with $\mathbb{T}^{-1} = \begin{pmatrix} \frac{3}{7} & -\frac{1}{7} \\ \frac{1}{7} & \frac{2}{7} \end{pmatrix}$, we obtain $\mathbb{U} = \mathbb{X} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ or $|u\rangle^x = |x_1\rangle + |x_2\rangle$. This small example demonstrates that we only have to know the transformation matrix between two bases to calculate the coefficients of a vector given in one basis in the other one.

1.5 Normalisation and Orthogonalisation of Vectors

A vector is called normalised, if the scalar product with itself equals 1, and when the scalar product of two different vectors yields zero, they are called orthogonal to each other. If a set of vectors fulfils the condition $\langle e_i|e_j\rangle = \delta_{ij}$, they form an orthonormalised or unitary vector system, respectively. To normalise a vector $|u\rangle$, one calculates $\frac{|u\rangle}{\sqrt{\langle u|u\rangle}} = \frac{|u\rangle}{|u|} = |e_u\rangle$, where $|e_u\rangle$ results as a normalised or 'unit'

vector in the direction of $|u\rangle$. Orthogonalisation of basis vectors can be achieved by a stepwise procedure ('Schmidt Orthogonalisation'), but there is a very convenient way to orthonormalise a complete basis set at once applying the *Löwdin Orthonormalisation*.

Let us note the basis $\{x_i\}$ as matrix $\mathbb{X}$. $\langle \mathbb{X}|\mathbb{X}\rangle = \mathbf{\Lambda}$ and is called the metrix of $\mathbb{X}$. The metrix is defined by:

$$\mathbf{\Lambda} = \begin{pmatrix} \langle x_1|x_1\rangle & \langle x_1|x_2\rangle & \cdots & \langle x_1|x_n\rangle \\ \langle x_2|x_1\rangle & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \langle x_n|x_1\rangle & \cdots & \cdots & \langle x_n|x_n\rangle \end{pmatrix}$$

We now form a new basis matrix $\mathbb{X}'$ by multiplying the original basis by $\mathbf{\Lambda}^{-\frac{1}{2}}$:

$$\mathbb{X} \cdot \mathbf{\Lambda}^{-\frac{1}{2}} = \mathbb{X}'$$

It is easily shown that the new basis set $\{x_i'\}$ consists of orthonormalised vectors:

$$\begin{aligned} \langle \mathbb{X}'|\mathbb{X}'\rangle &= \langle \mathbf{\Lambda}^{-\frac{1}{2}} \cdot \mathbb{X}|\mathbb{X} \cdot \mathbf{\Lambda}^{-\frac{1}{2}}\rangle = \mathbf{\Lambda}^{-\frac{1}{2}}\langle \mathbb{X}|\mathbb{X}\rangle\mathbf{\Lambda}^{-\frac{1}{2}} = \mathbf{\Lambda}^{-\frac{1}{2}} \cdot \mathbf{\Lambda} \cdot \mathbf{\Lambda}^{-\frac{1}{2}} \\ &= \mathbf{\Lambda}^{+\frac{1}{2}} \cdot \mathbf{\Lambda}^{-\frac{1}{2}} = \mathbb{1} \end{aligned}$$

It will prove very convenient to use orthonormalised basis sets in many cases. With the help of the Löwdin procedure, we can always convert any primary choice of basis set into an orthonormalised one.

1.6 Matrix Representation of the Scalar Product

The scalar product of two vectors is given by

$$\langle u|v\rangle = \sum_i u_i^* \cdot v_i = \mathbb{U}^+ \cdot \mathbb{V}$$

The latter expression corresponds to the matrix notation, and it is evident that the matrices $\mathbb{U}$ and $\mathbb{V}$ have to correspond to the same basis $\{e_i\}$. Explicitly, the vectors are given as:

$$\begin{aligned} |u\rangle &= c_{u_1} \cdot |e_1\rangle + c_{u_2} \cdot |e_2\rangle + \cdots + c_{u_n} \cdot |e_n\rangle = \mathbb{E} \cdot \mathbb{C}_u \\ |v\rangle &= c_{v_1} \cdot |e_1\rangle + c_{v_2} \cdot |e_2\rangle + \cdots + c_{v_n} \cdot |e_n\rangle = \mathbb{E} \cdot \mathbb{C}_v \end{aligned}$$

where $\mathbb{E}$ represents the basis, and the coefficients $(c_{u_1}, c_{u_2}, \ldots, c_{u_n})$ of $|u\rangle$ form the column vector $\mathbb{C}_u$, the coefficients of $|v\rangle$ the column vector $\mathbb{C}_v$. Thus, the

scalar product of $|u\rangle$ and $|v\rangle$ is given by:

$$\langle u|v\rangle = \mathbb{U}^+ \cdot \mathbb{V} = \mathbb{C}_u^+ \langle \mathbb{E}|\mathbb{E}\rangle \mathbb{C}_v = \mathbb{C}_u^+ \mathbf{\Lambda} \mathbb{C}_v$$

or simply by $\mathbb{C}_u^+ \mathbb{C}_v$, if the basis is orthonormalised.

The same formalism is valid in function space. Instead of $\{e_i\}$ we use $\{\varphi_i\}$

$$\psi_u = \mathbf{\Phi} \cdot \mathbb{C}_u$$
$$\psi_v = \mathbf{\Phi} \cdot \mathbb{C}_v$$

The scalar product $\langle \psi_u|\psi_v\rangle$ – corresponding to integration – is thus given as $\mathbb{C}_u^+ \langle \mathbf{\Phi}|\mathbf{\Phi}\rangle \mathbb{C}_v$ with $\langle \mathbf{\Phi}|\mathbf{\Phi}\rangle = \mathbb{S}$, the metrix, which is called the overlap integral matrix in function space.

$$\mathbb{S} = \begin{pmatrix} \langle \varphi_1|\varphi_1\rangle & \cdots & \langle \varphi_1|\varphi_n\rangle \\ \cdots & & \\ \langle \varphi_n|\varphi_1\rangle & \cdots & \langle \varphi_n|\varphi_n\rangle \end{pmatrix}$$

$$\mathbb{S} = \begin{pmatrix} \int \varphi_1^* \varphi_1 \, d\tau & \cdots & \int \varphi_1^* \varphi_n \, d\tau \\ \cdots & & \\ \int \varphi_n^* \varphi_1 \, d\tau & \cdots & \int \varphi_n^* \varphi_n \, d\tau \end{pmatrix}$$

if $\langle \varphi_i|\varphi_j\rangle = \delta_{ij}$, then $\mathbb{S} = \mathbb{1}$.

The Löwdin orthonormalisation can be analogously performed in function space: starting from an arbitrary basis $\{\varphi_i\}$ one obtains $\langle \mathbf{\Phi}|\mathbf{\Phi}\rangle = \mathbb{S}$, and transforming the basis by $\mathbf{\Phi}' = \mathbf{\Phi} \cdot \mathbb{S}^{-\frac{1}{2}}$, one has $\langle \mathbf{\Phi}'|\mathbf{\Phi}'\rangle = \mathbb{1}$, which means that $\mathbf{\Phi}'$ is the desired orthonormalised basis.

1.7 Dual Vector Space and Hilbert Space

While talking of scalar products, we have silently assumed that we can find a transposed, complex conjugated form for every vector as well, thus implying the condition:

$$\exists \langle a|\forall|a\rangle \in V^n$$

A vector space fulfilling this condition is called a *Dual Vector Space*. This condition is essential for a vector space describing a physical system, as the evaluation of physical quantities implies the formation of scalar products. In addition to that, two further conditions will be imposed on the vector space we are going to use for the description of matter, namely the two *Cauchy convergence criteria*:

$$\sum_{i=1}^{\infty} \langle \Psi_i|\Psi_i\rangle < \infty$$

and

$$\lim_{M,N\to\infty} \sum_n |\Psi_n^M - \Psi_n^N|^2 = \emptyset$$

employing different representations of Ψ.

These criteria are also termed 'quadratic convergence' of the vector space/function space, and a dual vector space fulfilling them is called a *Hilbert Space*. In the next section we will see that quadratic convergence is required in order to guarantee the normalisability of a system to a finite number of particles.

1.8 Probability Concept and the Ψ Function

The probability concept implies that we can describe any physical system by a function/vector containing all the information about this system, i.e. its properties as a function of coordinates, and that we can evaluate the probability for any state of the system from this function by the product $\Psi^*\Psi$. The overall probability is then given by the scalar product

$$\int_{-\infty}^{\infty} \Psi^* \cdot \Psi \, d\tau = \langle \Psi | \Psi \rangle$$

It is now understood, why quadratic convergence of the vector space to which $|\Psi\rangle$ belongs is required, as otherwise the overall probability could not be restrained to a finite value.

The probability function is usually formulated as a function of space coordinates and the time coordinate as $\Psi(r_i, t)$, but in some cases the use of momentum coordinates ('momentum space') and time as $\Psi(|p_i\rangle, t)$ is advantageous.

The concept describing a physical system by a probability function/vector depending on a few variables implies that other physical variables must be evaluated from this function/vector. The mathematical instruments achieving this are called *operators*, and the general definition and properties of such operators will be detailed in the next section.

1.9 Operators

In this textbook, an operator is marked with a 'hat', (e.g., $\hat{H}$). The operator's action in vector space can be illustrated, if we consider the vector $|x\rangle$ and an operator $\hat{A}$: $\hat{A}|x\rangle = |x'\rangle$, i.e., the operator transforms the original vector, in general by changing its amount and its direction, into a new vector. One also says that the original vector is mapped onto an image vector. Some examples of operators are

$\widehat{C_2}$, representing a rotation by 180 degrees and $\widehat{C_\varphi}$, corresponding to a rotation by the angle φ. Another example is the inversion operator $\hat{i}$, by which every component of the vector is mapped onto its inverse, such as $x \to -x$. In V^3 the matrix $\mathbb{I}$ representing this operator is:

$$\mathbb{I} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

As the inversion operator and this matrix are associated by $\hat{i}\mathbb{X} = \mathbb{X}\mathbb{I}$, every operator can be represented in matrix form, once a vector space has been defined by a concrete basis.

It will be useful to summarise a few basic definitions and rules for operators. There are two types of operator: (i) *regular* operators, for which a one-to-one correspondence between original vectors and image vectors exists; and (ii) *singular* operators, for which the correspondence is only unidirectional from original vectors to image vectors. The previously given inversion operator is an example of a regular operator, the square operator $(\ )^2$ is a singular operator, as it is not possible to distinguish from a value 4 whether the original value in the domain of the operator was 2 or -2 (the set of all original vectors is called its 'domain').

An *inverse* operator multiplied with its source operator results in the *unity* operator, according to $\hat{A}^{-1}\hat{A} = \hat{1}$.

In general, operators are not commutative, i.e., $\hat{A}\hat{B} \neq \hat{B}\hat{A}$.

Linear operators are defined by the following conditions:

- $\hat{A}(|u\rangle + |v\rangle) = \hat{A}|u\rangle + \hat{A}|v\rangle$
- $\hat{A}(k|u\rangle) = k\hat{A}|u\rangle$, k being a scalar
- $(\widehat{A_1} + \widehat{A_2})|u\rangle = \widehat{A_1}|u\rangle + \widehat{A_2}|u\rangle$

Most of the operators we are using in vector space theory of matter are linear operators.

1.10 Representation of Operators in a Basis

In a given basis $\{e_i\}^n$, an operator will act on each of the basis vectors by transforming it into another vector, which will also be an element of the same vector space $\hat{A}|e_1\rangle = |e'\rangle \in V^n$, and can be constructed, therefore, as a linear combination of all original basis vectors:

$$\hat{A}|e_1\rangle = a_{11}|e_1\rangle + a_{12}|e_2\rangle + \cdots + a_{1n}|e_n\rangle = \sum_i a_{1i}|e_i\rangle$$

$$\hat{A}|e_2\rangle = \sum_i a_{2i}|e_i\rangle, \quad \text{until } \hat{A}|e_n\rangle = \sum_i a_{ni}|e_i\rangle$$

The first index of $a_{1i}, a_{2i}, \ldots a_{ni}$ specifies which basis vector is transformed. All of these equations can be shortly summarised in matrix form as

$$\hat{A}\mathbb{E} = \mathbb{E}\mathbb{A}^e \tag{1}$$

where $\mathbb{A}^e$ contains all coefficients a_{ij} and is thus a complete representation of the operator in the basis $\{e_i\}^n$. If Eq. (1) is multiplied with the basis in the bra space, one obtains: $\langle \mathbb{E}|\hat{A}\mathbb{E}\rangle = \langle \mathbb{E}|\mathbb{E}\rangle\mathbb{A}^e$ and from this

$$\mathbb{A}^e = \frac{\langle \mathbb{E}|\hat{A}\mathbb{E}\rangle}{\langle \mathbb{E}|\mathbb{E}\rangle}$$

which is the general expression for the representation of the operator in the given basis. This expression is also termed the *expectation value* of the operator $\hat{A}$. For orthonormal bases, this reduces to $\mathbb{A}^e = \langle \mathbb{E}|\hat{A}\mathbb{E}\rangle$.

The matrix $\mathbb{A}^e$ contains the following elements:

$$\mathbb{A}^e = \begin{pmatrix} \langle e_1|\hat{A}e_1\rangle & \langle e_1|\hat{A}e_2\rangle & \cdots & \langle e_1|\hat{A}e_n\rangle \\ \langle e_2|\hat{A}e_1\rangle & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \langle e_n|\hat{A}e_1\rangle & \cdots & \cdots & \langle e_n|\hat{A}e_n\rangle \end{pmatrix}$$

In function space the analogous representation of an operator in a basis of functions is given as

$$\mathbb{A}^\varphi = \frac{\langle \mathbf{\Phi}|\hat{A}\mathbf{\Phi}\rangle}{\langle \mathbf{\Phi}|\mathbf{\Phi}\rangle} \quad \text{or} \quad \langle \mathbf{\Phi}|\hat{A}\mathbf{\Phi}\rangle$$

for an orthonormalised basis set. In detail, this matrix consists of the elements:

$$\mathbb{A}^\varphi = \begin{pmatrix} \int \varphi_1^* \hat{A}\varphi_1 \, d\tau & \cdots & \int \varphi_1^* \hat{A}\varphi_n \, d\tau \\ \cdots & \cdots & \cdots \\ \int \varphi_n^* \hat{A}\varphi_1 \, d\tau & \cdots & \int \varphi_n^* \hat{A}\varphi_n \, d\tau \end{pmatrix}$$

In order to recall the correspondence between scalar products and integrals, the elements of this matrix have been written as integrals.

1.11 Change of Basis in Representations of Operators

If we have two different bases $\{e_i\}$ and $\{f_i\}$, whose connection is given by $\mathbb{F} = \mathbb{E}\mathbb{T}$ (linear transformation), we can connotate the equation $\hat{A}|u\rangle = |v\rangle$ as $\mathbb{V}^e = \mathbb{A}^e\mathbb{U}^e$ or $\mathbb{V}^f = \mathbb{A}^f\mathbb{U}^f$, depending on the chosen basis.

Based on $\mathbb{U}^e = \mathbb{T}\mathbb{U}^f$ and $\mathbb{V}^e = \mathbb{T}\mathbb{V}^f$ we can formulate $\mathbb{T}\mathbb{V}^f = \mathbb{A}^e\mathbb{T}\mathbb{U}^f$, which upon multiplication with $\mathbb{T}^{-1}$ from the left leads to

$$\mathbb{V}^f = \mathbb{T}^{-1}\mathbb{A}^e\mathbb{T}\mathbb{U}^f = \mathbb{A}^f\mathbb{V}^f,$$

from which follows

$$\mathbb{A}^f = \mathbb{T}^{-1}\mathbb{A}^e\mathbb{T}$$

Such a transformation is called *similarity transformation*, and if $\mathbb{T}^{-1} = \mathbb{T}^+$, *unitary transformation*.

Changing the basis (e.g., a coordinate system) must not change the physics of the system. Besides the expectation value, two further properties of a matrix are invariant in similarity transformations, namely the determinant and the trace of the matrix:

$$det(\mathbb{A}) = det(\mathbb{T}^{-1}\mathbb{A}\mathbb{T})$$

$$tr(\mathbb{A}) = \sum_i a_{ii} = tr(\mathbb{T}^{-1}\mathbb{A}\mathbb{T})$$

Consequently, all physical quantities associated with the operator will have to be associated with its expectation value, the determinant, and/or trace of its representation matrix.

Test Questions Related to this Chapter

1. Why can we treat functions as vectors?
2. What are the differences between a model and a theory?
3. Why does vector space designed to describe matter have to be a Hilbert space?
4. Can we use the probability concept in classical physics?
5. What possibilities exceeding classical physics offer the probability concept in the description of particle behaviour?
6. What consequences does the probability concept have for the evaluation of physical variables?
7. What is a unitary transformation, and which characteristics of a matrix are invariant to it?