CONCEPTS OF CONVERGENCE IN MATHEMATICAL CHEMISTRY

Per-Olov LÖWDIN
Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, FL 32611, USA

Abstract

Some concepts of convergence used in mathematical chemistry are briefly reviewed: number convergence, uniform and non-uniform function convergence, convergence in norm and binary product, operator convergence, computer convergence, etc. Some properties of the abstract Hilbert space and some of its realizations in mathematical chemistry are discussed. Finally, it is pointed out that the scattering wave functions of importance in the theory of chemical reactions are limit points of the L^2 Hilbert space – not in the norm, but in the sense of a non-uniform point-by-point convergence, which is of essential value in practical applications.

1. Introduction

In various parts of mathematical chemistry, one often deals with infinite sequences of numbers, objects, etc., and their limits, and it is evident that the concept of convergence is fundamental in this connection. In their work, most theoretical chemists take these convergence properties for granted, and they are seldom discussed. The remarkable fact is that quite a few different convergence concepts are used in various connections, and some of the most important ones will be briefly reviewed in this paper.

1.1. NUMBER CONVERGENCE

Let us consider an infinite sequence of real or complex numbers a_1, a_2, a_3, $a_4, \ldots, a_n, \ldots$. Such a sequence is said to have a limit a, provided that for every positive number ε – however small – there exists a mapping $\varepsilon \rightarrow N(\varepsilon)$ such that, whenever $n > N(\varepsilon)$, one has for the absolute value $|\,|$ of the difference:

$$|a_n - a| < \varepsilon, \qquad n > N(\varepsilon). \tag{1}$$

In pure mathematics, one proves the existence of convergence by giving at least a rough estimate of the function $N(\varepsilon)$. Since one usually does not know the limit a in advance, relation (1) is often replaced by Cauchy's necessary and sufficient criterion for convergence:

*Professor Emeritus, Uppsala University, Uppsala, Sweden.

$$|a_m - a_n| < \varepsilon, \qquad m > n > N(\varepsilon), \tag{2}$$

which was formulated in the first half of the 18th century. In 1872, it was pointed out by Cantor and Dedekind that *irrational numbers* had not been properly defined in mathematics, and that so-called Cauchy sequences of rational numbers satisfying (2) could be used for this purpose. A simple example of number convergence is given by the relation

$$\lim_{n \to \infty} \sum_{k=1}^{n} (-1)^k (1/k) = \log 2. \tag{3}$$

1.2. FUNCTIONAL CONVERGENCE

Let us next consider a series of real or complex functions $f_1(x)$, $f_2(x)$, $f_3(x), \ldots, f_n(x) \ldots$ of the variable x. Such a sequence is said to have a limit function $f(x)$, provided that there exists a function $N(\varepsilon, x)$ such that

$$|f_n(x) - f(x)| < \varepsilon, \qquad n > N(\varepsilon, x). \tag{4a}$$

Since the function N depends also on the value of the variable x, one speaks of *point-by-point convergence*. If the function N is independent of the value of x in a certain interval $[a, b]$, one speaks of *uniform convergence* in that interval:

$$|f_n(x) - f(x)| < \varepsilon, \qquad n > N(\varepsilon). \tag{4b}$$

2. Some properties of the abstract Hilbert space and its realizations

2.1. THE AXIOMS OF THE ABSTRACT HILBERT SPACE

When modern quantum theory was developed in 1925–26 [1], it was given three completely different formulations. In his wave mechanics, Schrödinger interpreted the basic physical observables x and p as *operators*, whereas in the matrix mechanics developed by Heisenberg, Born, and Jordan, they were interpreted as *matrices*. In Dirac's abstract formulation, they were finally considered as content-less non-commutative quantities called *q-numbers*. The only thing common for the three formulations seemed to be the commutation relation $px - xp = h/2\pi i$. The three different approaches were unified by von Neumann [2] by introducing the concept of the *abstract Hilbert space*, which has turned out to be one of the most important tools in modern quantum theory.

An abstract Hilbert space is a linear space $A = \{f\}$ having a binary product $\langle f | g \rangle$ which satisfies the following four axioms:

axiom 1: $\langle f | g_1 + g_2 \rangle = \langle f | g_1 \rangle + \langle f | g_2 \rangle,$ (5)

axiom 2: $\langle f | g \cdot \alpha \rangle = \langle f | g \rangle \alpha,$ (6)

axiom 3: $\langle g | f \rangle = \langle f | g \rangle^{\star},$ (7)

axiom 4: $\langle f | f \rangle \geq 0,$ and $=0$ if and only if $f = 0.$ (8)

The first two axioms indicate that the binary product is linear in the second position, the third that it is Hermitian symmetric, and the fourth that the space is positive definite. Since one has the property $\langle f \cdot \alpha | g \rangle = \alpha^{\star} \langle f | g \rangle$, one sometimes says that the first position is anti-linear. It is further convenient to introduce the *norm* or length $\| f \|$ of an element f by the relation $\| f \| = \langle f | f \rangle^{1/2}$. By using the axioms, one immediately obtains Schwarz's inequality and the related triangular inequality:

$$|\langle f | g \rangle| \leq \| f \| \cdot \| g \|, \tag{9}$$

$$\big| \| f \| - \| g \| \big| \leq \| f + g \| \leq \| f \| + \| g \|. \tag{10}$$

2.2. CONVERGENCE IN THE NORM

An infinite series of elements $f_1, f_2, f_3, \ldots, f_n, \ldots$ in the linear space $A = \{ f \}$ is said to be *convergent in the norm* toward the limit element f if

$$\| f_n - f \| < \varepsilon, \qquad \text{whenever } n > N(\varepsilon). \tag{11}$$

The series is said to be a *Cauchy sequence*, if $\| f_m - f_n \| < \varepsilon$, whenever $m > n > N(\varepsilon)$, and every such sequence may be used to define the limit element f. In addition, one also speaks of *binary product convergence* of the sequence $\{f_n\}$ in the case where

$$|\langle f | g \rangle - \langle f_n | g \rangle| < \varepsilon, \qquad \text{whenever } n > N(\varepsilon, g) \tag{12}$$

for all elements g, and this is obviously a type of number convergence. Convergence in the norm is sometimes referred to as "strong convergence", whereas convergence in the binary product is termed "weak convergence". Due to the relation

$$|\langle f | g \rangle - \langle f_n | g \rangle| < |\langle f - f_n | g \rangle| < \| f - f_n \| \cdot \| g \|, \tag{13}$$

one sees immediately that strong convergence implies weak convergence, whereas a simple example shows that the reverse is not true.

According to von Neumann, a linear space $A = \{ f \}$ with a binary product satisfying axioms 1–4 becomes an *abstract Hilbert space* if it satisfies two more axioms:

axiom 5: It contains all its *limit points* in the norm.

axiom 6: It is separable in the sense that it contains at least one enumerable set $\{h_1, h_2, h_3, \ldots, h_n, \ldots\}$ which is *everywhere dense* in $A = \{f\}$, so that for every f there is at least one h_k so that $\|f - h_k\| < \varepsilon$.

Some authors have claimed that the last axiom is of only marginal importance in theoretical physics and mathematical chemistry, but in reality it is of major value, also for the applications. A set $B = \{\varphi\}$ is said to be *orthonormal* if it satisfies the two conditions $\|\varphi\| = 1$ and $\langle\varphi|\varphi'\rangle = 0$, and axiom 6 shows that any such set must necessarily be enumerable. If one starts out from the enumerable set $\{h_1, h_2, h_3, \ldots, h_n, \ldots\}$, and applies Schmidt's successive orthonormalization procedure, one obtains an orthonormal set $B = \{\varphi_k\}$ which is *complete* in the sense that, if $\langle g|\varphi_k\rangle = 0$ for all k, then one must have $g = 0$, i.e. there are no functions $g \neq 0$ outside the set that are orthogonal to the functions in the set. Hence, axiom 6 ensures the existence of at least one complete orthonormal set $\varphi = \{\varphi_k\}$.

Next, we will show that the set $\varphi = \{\varphi_k\}$ *spans* the abstract Hilbert space in the sense that there exists an *expansion theorem* of the form

$$f = \lim_{n\to\infty} \sum_{k=1}^{n} \varphi_k\langle\varphi_k|f\rangle = \sum_{k=1}^{\infty} \varphi_k\langle\varphi_k|f\rangle, \tag{14}$$

which is convergent in the *norm*. For this purpose, we will consider the sequence

$$r_n = f - \sum_{k=1}^{n} \varphi_k\langle\varphi_k|f\rangle, \tag{15}$$

for $n = 1, 2, 3, \ldots$. Using the orthonormality property of the set $\varphi = \{\varphi_k\}$, one obtains immediately

$$\|r_n\|^2 = \langle f|f\rangle - \sum_{k=1}^{n} |\langle\varphi_k|f\rangle|^2 \geq 0, \tag{16}$$

which is the famous *Bessel inequality*, which indicates that the series

$$\sum_{k=1}^{\infty} |\langle\varphi_k|f\rangle|^2 \tag{17}$$

is always number convergent. Hence, one has

$$\|r_m - r_n\|^2 = \sum_{k=n+1}^{m} |\langle\varphi_k|f\rangle|^2 < \varepsilon, \quad \text{whenever } m > n > N(\varepsilon), \tag{18}$$

which implies that $\{r_n\}$ is a Cauchy sequence which uniquely defines a limit point in the norm r:

$$r = f - \sum_{k=1}^{\infty} \varphi_k \langle \varphi_k | f \rangle. \tag{19}$$

However, since $\langle r_n | \varphi_k \rangle = 0$ for all k, and weak convergence follows from strong convergence, one also has $\langle r | \varphi_k \rangle = 0$ for all k, which means that $r = 0$ as a consequence of the completeness of the orthonormal set $\varphi = \{\varphi_k\}$. Hence, the expansion theorem (14) is proven. Letting n go to infinity in relation (16) one also obtains

$$\langle f | f \rangle = \sum_{k=1}^{\infty} |\langle \varphi_k | f \rangle|^2, \tag{20}$$

which is *Parseval's first relation*, which is number convergent. We note that the relations (14) and (20) each may be used as alternative definitions of the concept of *completeness* for orthonormal sets.

2.3. MAPPINGS AND OPERATORS

If $A = \{f\}$ and $B = \{g\}$ are two sets in general, one says that any pairing of the elements $[f, g]$ defines a pairwise *relation*. If, for a given f, the second element g is unique, one often speaks of a *mapping* $f \to g$, and one says that g is the *image* of the element f. In the case where one has $f \leftrightarrow g$, one speaks of a *one-to-one mapping*, and we note that, since the time of Cantor, this idea forms the basis for the concept of *cardinality* and the set theory of the integers 0, 1, 2, 3, . . . , and the transfinite numbers $\aleph_0, \aleph_1, \ldots$, etc. One frequently says that a mapping $f \to g$ is described by an *operator* T, and one writes $Tf = g$.

If $\{f\}$ is a linear space, one says that T is a *linear operator*, if it preserves the linearity of the space, so that $T(f_1 \cdot \alpha_1 + f_2 \cdot \alpha_2) = T(f_1) \cdot \alpha_1 + T(f_2) \cdot \alpha_2$. If the space $\{f\}$ has a binary product $\langle f | g \rangle$, which satisfies axioms 1–4, one says that T has the *adjoint operator* $T^\dagger$ defined through the relation

$$\langle f | Tg \rangle = \langle T^\dagger f | g \rangle, \tag{21}$$

and one obtains immediately the rules $(T_1 \cdot \alpha_1 + T_2 \cdot \alpha_2)^\dagger = (T_1)^\dagger \cdot \alpha_1^* + (T_2)^\dagger \cdot \alpha_2^*$, $(T_1 \cdot T_2)^\dagger = (T_2)^\dagger (T_1)^\dagger$. Finally, if $A = \{f\}$ is an abstract Hilbert space, one says that the operator T is defined on the *domain* $D(T)$ consisting of the elements f, if both f and its image Tf are elements of the Hilbert space.

Let us now consider some simple examples of operators working on the abstract Hilbert space. We note that Dirac considered the *bracket* $\langle a | b \rangle$ as the scalar product of a *bra-vector* $\langle a |$ and a *ket-vector* $| b \rangle$, and he then defined the *ket-bra operator* $T = | b \rangle \langle a |$ through the relations

$$T = |b\rangle\langle a|, \quad Tf = b\,\langle a|f\rangle. \tag{22}$$

Using the definition, one obtains immediately the theorems

$$T^{\dagger} = |a\rangle\langle b|, \quad T^2 = \langle a|b\rangle\, T, \quad \mathrm{Tr}\, T = \langle a|b\rangle, \tag{23}$$

where the symbol Tr (=Trace) indicates the sum over all the eigenvalues of the operator T, which has a single non-vanishing eigenvalue $\lambda = \langle a|b\rangle$, which is easily seen to be non-degenerate. The use of brackets (,) to denote binary products was introduced by the mathematicians long before Dirac, but we note that they preferred to have the first position linear; hence, one has the connection $(b, a) \equiv \langle a|b\rangle$. Instead of Dirac's ket-bra operator $|b\rangle\langle a|$, they used the dyadic product operator with the notation $(.,a)b$. In reading the mathematical literature, it is therefore worthwhile to remember the connection formulas

$$(b, a) \equiv \langle a|b\rangle, \quad (.,a)b \equiv |b\rangle\langle a|. \tag{24}$$

2.4. OPERATOR CONVERGENCE

Let us now consider the space $B = \{T\}$ formed by all linear operators T defined on the abstract Hilbert space $A = \{f\}$. Let us further consider an infinite sequence $T_1, T_2, T_3, \ldots, T_n, \ldots$ of such operators. Such a sequence is said to have a limit element T, if

$$\|(T - T_n)f\| < \varepsilon, \qquad \text{whenever } n > N(\varepsilon, f), \tag{25}$$

and it is obvious that such an *operator convergence* can be of a point-by-point type for the individual elements f or uniform, when $N(\varepsilon, f)$ does not depend on f. It should be emphasized that, in mathematics, there are many more definitions of the concept of "operator convergence", but also that the definition (25) is the most convenient one for our present purposes.

In the field of mathematical chemistry, one says that an operator O is an *orthogonal projector* of order g, if it satisfies the relations

$$O^2 = O, \quad O^{\dagger} = O, \quad \mathrm{Tr}\, O = g. \tag{26}$$

In such a case, the operator $P = 1 - O$ is said to be the projector for its *orthogonal complement*, and it satisfies the relations

$$P^2 = P, \quad P^{\dagger} = P, \quad \mathrm{Tr}\, P = \infty. \tag{27}$$

We note that the relation $1 = O + P$ forms a trivial example of a "resolution of the identity operator 1". A projector having $g = 1$ is said to be a *primitive projector*, and a simple example is given by the ket-bra operator $O_k = |\varphi_k\rangle\langle\varphi_k|$. Next, we will consider the sum of n such primitive operators:

$$Q_n = \sum_{k=1}^{n} O_k = \sum_{k=1}^{n} |\varphi_k\rangle\langle\varphi_k| . \tag{28}$$

Using the orthonormality property of the set $\varphi = \{\varphi_k\}$, one obtains directly the relations

$$Q_n^2 = Q_n, \quad Q_n^\dagger = Q_n, \quad \mathrm{Tr}\, Q_n = n, \tag{29}$$

and Q is hence an orthogonal projector of order n. We will further consider the projector $P_n = 1 - Q_n$ for its orthogonal complement, which satisfies the relations

$$P_n^2 = P_n, \quad P_n^\dagger = P_n, \quad \mathrm{Tr}\, P_n = \infty. \tag{30}$$

Let us now once more consider the remainder r_n in the expansion theorem, which is defined by relation (15). One obtains directly

$$r_n = f - \sum_{k=1}^{n} \varphi_k \langle\varphi_k|f\rangle = f - \sum_{k=1}^{n} O_k f = (1 - \sum_{k=1}^{n} O_k) f = (1 - Q_n) f = P_n f, \tag{31}$$

and further

$$\begin{aligned} \| r_n \|^2 &= \|(1 - Q_n) f\|^2 = \|P_n f\|^2 = \langle P_n f | P_n f\rangle = \langle f | P_n^\dagger P_n | f\rangle \\ &= \langle f | P_n | f\rangle = \langle f | 1 - Q_n | f\rangle = \langle f|f\rangle - \sum_{k=1}^{n} \langle f|\varphi_k\rangle\langle\varphi_k|f\rangle \to 0, \end{aligned} \tag{32}$$

where, according to relation (20), the right-hand member goes to zero when n goes to infinity. Hence, one has the relation

$$\lim_{n\to\infty} \|(1 - Q_n) f\| = 0, \tag{33}$$

which implies that the operator sequence $Q_1, Q_2, Q_3, \ldots, Q_n, \ldots$ converges to the identity operator 1, and obtains the following *resolution of the identity*:

$$1 = \lim_{n\to\infty} Q_n = \sum_{k=1}^{\infty} |\varphi_k\rangle\langle\varphi_k|, \tag{34}$$

where the sequence $\{Q_n\}$ is *operator convergent*. It is evident that this is a necessary and sufficient condition for the *completeness* of the set $\varphi = \{\varphi_k\}$.

A convenient tool in mathematical chemistry is the use of boldface symbols for rectangular matrices: column vectors, row vectors, quadratic matrices, etc. If $\boldsymbol{A}$ and $\boldsymbol{B}$

are two rectangular matrices of order $m \times p$ and $p \times n$, then their product $\boldsymbol{C} = \boldsymbol{AB}$ is a matrix of order $m \times n$ having the elements $C_{kl} = \sum_\alpha A_{k\alpha} B_{\alpha l}$, where one has multiplied the rows of the first matrix in order with the columns of the second matrix. Using this short-hand, one can write the orthonormality property of the set $\varphi = \{\varphi_k\}$ and its completeness in terms of the two conditions:

$$\langle \varphi | \varphi \rangle = \mathbf{1}, \qquad 1 = |\varphi\rangle \langle \varphi|, \tag{35}$$

where $\mathbf{1}$ is the unit matrix. One may use the second relation (35) as a strict rule, if one observes that the operator convergence is changed into other types of convergence depending on the circumstances. One has, for example,

$$f = 1 \cdot f = |\varphi\rangle \langle \varphi | f \rangle = \sum_{k=1}^{\infty} \varphi_k \langle \varphi_k | f \rangle; \tag{36}$$

$$\langle f | f \rangle = \langle f | 1 | f \rangle = \langle f | \varphi \rangle \langle \varphi | f \rangle = \sum_{k=1}^{\infty} \langle f | \varphi_k \rangle \langle \varphi_k | f \rangle; \tag{37}$$

$$\langle f | g \rangle = \langle f | 1 | g \rangle = \langle f | \varphi \rangle \langle \varphi | g \rangle = \sum_{k=1}^{\infty} \langle f | \varphi_k \rangle \langle \varphi_k | g \rangle; \tag{38}$$

where relation (36) is convergent in the norm, whereas relations (37) and (38) are number convergent. Similarly, if f is an element in the domain $D(T)$ of the operator T, one obtains

$$\begin{aligned} Tf &= 1 \cdot T \cdot 1 \cdot f = |\varphi\rangle \langle \varphi | T | \varphi \rangle \langle \varphi | f = \varphi \langle \varphi | T | \varphi \rangle \langle \varphi | f \rangle \\ &= \varphi \boldsymbol{T} \langle \varphi | f \rangle = \sum_{k,l=1}^{\infty} \varphi_k T_{kl} \langle \varphi_l | f \rangle, \end{aligned} \tag{39}$$

which relation is convergent in the norm. Here, we have used the matrix notations

$$\boldsymbol{T} = \langle \varphi | T | \varphi \rangle, \quad T_{kl} = \langle \varphi_k | T | \varphi_l \rangle. \tag{40}$$

For the operator T itself, one further obtains

$$T = 1 \cdot T \cdot 1 = |\varphi\rangle \langle \varphi | T | \varphi \rangle \langle \varphi | = |\varphi\rangle \boldsymbol{T} \langle \varphi | = \sum_{k,l=1}^{\infty} |\varphi_k\rangle T_{kl} \langle \varphi_l |, \tag{41}$$

which relation is operator convergent. If one introduces the notation

$$P_{lk} = |\varphi_k\rangle \langle \varphi_l|, \tag{42}$$

one gets the expansion

$$T = \sum_{k,l=1}^{\infty} |\varphi_k\rangle T_{kl} \langle \varphi_l| = \sum_{k,l=1}^{\infty} T_{kl} P_{lk} \,, \tag{43}$$

and one realizes that every linear operator T is expandable in terms of its matrix elements and the fundamental operators P_{lk}, which apparently span the operator space $\{T\}$ and form some form of *basis* in this space. The operators P_{lk} are sometimes referred to as the *fundamental units* of the operator space, and we note that they satisfy the product formula

$$P_{lk} P_{nm} = \delta_{lm} P_{nk}, \tag{44}$$

which tells us – among other things – that the diagonal units are projectors of order 1, whereas the non-diagonal units are nil-potent of order 2. At this point, it is convenient to introduce a *binary product* in the operator space $B = \{T\}$ through the relation

$$\{T_1 | T_2\} = \mathrm{Tr} T_1^{\dagger} T_2 = \sum_{k=1}^{\infty} \langle \varphi_k | T_1 | \varphi_l \rangle^* \langle \varphi_k | T_2 | \varphi_l \rangle, \tag{45}$$

which is of *Hilbert–Schmidt type*, and it is then easily verified that the fundamental units $\{P_{lk}\}$ form an orthonormal basis in the operator space, which is also complete. The operators T having a finite norm $\|T\| = \{T|T\}^{1/2}$ form again an abstract Hilbert space, which will be referred to as the *abstract Hilbert–Schmidt space*.

The operators T are *mappings* of the elements in the space $A = \{f\}$, which is often referred to as the *carrier space* for these operators. In the same way, one may consider the mappings $\hat{M}$ of the operators T, which are sometimes referred to as *super-operators*. The superoperators $\hat{M}$ have the operator space $\{T\}$ as their carrier space, and of particular importance in mathematical chemistry is the Liouvillian super-operator $\hat{L}$ and the super-evolution operator $\hat{S}$, which are defined through the relations:

$$\hat{L}T = HT - TH, \quad \hat{S}T = STS^{-1}, \tag{46}$$

where H is the ordinary Hamiltonian and S is the conventional evolution operator $S = \exp\{(-2\pi i/h)H(t-t_0)\}$. For further details as to the theory of superoperators and their use in mathematical chemistry, the reader is referred elsewhere [3].

We note that all formulas (35)–(44) are valid in the abstract Hilbert space and that we – so far – have not defined the binary product $\langle f|g\rangle$ we are using.

2.5. REALIZATIONS OF THE ABSTRACT HILBERT SPACE

A characteristic feature of modern mathematics is that one tries to develop deductive theories, which are as *content-less* as at all possible, by using a number of undefined quantities which later may be given different types of *realizations* giving rise to different *models* of the abstract theory. In the theory of the abstract Hilbert space

$A = \{f\}$, the main undefined quantities are the elements f of the space and the binary product $\langle f | g \rangle$, and we will now consider some realizations of these concepts, which are of essential importance in mathematical chemistry.

The oldest realization is the *sequential Hilbert space* $\mathcal{H}_0$, in which the elements f are infinite column vectors $\boldsymbol{c} = \{c_k\}$ with the property that

$$\sum_{k=1}^{\infty} |c_k|^2 < \infty. \tag{47}$$

In this space, one starts by defining the binary product through the relation

$$\langle \boldsymbol{c} | \boldsymbol{d} \rangle = \boldsymbol{c}^\dagger \boldsymbol{d} = \sum_{k=1}^{\infty} c_k^* d_k, \tag{48}$$

and this means that one has now to show that axioms 1–6 for the abstract Hilbert space are now satisfied as *theorems*. Once this is accomplished, one can conclude that all other theorems in the theory of the abstract Hilbert space are automatically valid.

In another important realization of $A = \{f\}$, the elements f are *wave functions* $\Psi = \Psi(X)$ for a many-particle system with the composite coordinate $X = (x_1, x_2, x_3, \ldots, x_N)$, where each x_k is a combined space–spin coordinate $x_k = (\boldsymbol{r}_k, \zeta_k)$, which are absolutely quadratically integrable, so that

$$\int |\Psi(X)|^2 \, \mathrm{d}X < \infty, \tag{49}$$

where the meaning of the integration sign $\int \mathrm{d}X$ is that one integrates over all the $3N$ space coordinates $\boldsymbol{r}_k$ and sumes over the N spin coordinates ζ_k. In this case, one starts by defining the binary product through the formula

$$\langle \Psi_1 | \Psi_2 \rangle = \int \Psi_1(X)^* \Psi_2(X) \mathrm{d}X. \tag{50}$$

It is clear that axioms 1–4 are immediately valid as theorems, but that axioms 5 and 6 will require a more careful consideration. In fact, if the integration is defined as a Riemann integration, they are usually not valid, whereas it may be shown that they are true if one instead uses the more general integral concept introduced by *Lebesgue* in the early 1900's. Since the Lebesgue integration is essential, one often refers to this realization of $A = \{f\}$ as the *L^2 Hilbert space.*

In this model, one may also use a continuous representation of certain operators T by means of their *kernels* $T(X|X')$, introduced through the relation:

$$T\Psi(X) = \int T(X|X')\Psi(X')\mathrm{d}X'. \tag{51}$$

That such a representation exists follows from formula (39), which in this case takes the form

$$T\Psi = \sum_{k,l=1}^{\infty} \varphi_k T_{kl} \langle \varphi_l | \Psi \rangle = \sum_{k,l=1}^{\infty} \varphi_k T_{kl} \int \varphi_l(X')^* \Psi(X') \mathrm{d}X ; \tag{52}$$

hence, one has for the associated kernel

$$T(X|X') = \sum_{k,l=1}^{\infty} \varphi_k(X) T_{kl} \varphi_l(X')^*, \tag{53}$$

and it may then be shown that, for certain types of operators, the series is point-by-point convergent in the (X, X')-space. The identity operator 1 is represented by the unit matrix **1**, and – in order to provide it with a kernel – Dirac introduced his famous δ-function $\delta(X - X')$, so that one obtains

$$\sum_{k=1}^{\infty} \varphi_k(X) \varphi_k(X')^* = \delta(X - X'). \tag{54}$$

This series is not convergent in the sense of any of the definitions given above, and – in order to give it a strict meaning – one has to go over to the theory of the *distribution functions*. It is claimed that Dirac's δ-function is of particular importance in the treatment of scattering problems, e.g. in the theory of chemical reactions, and that it is necessary in order to obtain a reasonable "normalization" of the eigenfunctions in the continuum. For this reason, we will now briefly review the underlying eigenvalue problem.

2.6. THE BOUNDARY CONDITIONS IN THE TREATMENT OF THE EIGENVALUE PROBLEM OF THE HAMILTONIAN H

In Schrödinger's wave mechanics, the time-independent equation for the stationary states has the form

$$H\Psi = E\Psi, \tag{55}$$

where H is the classical Hamiltonian with the momentum vector p_k replaced by the operator $(h/2\pi i)(\partial/\partial x_k, \partial/\partial y_k, \partial/\partial z_k)$. Equation (55) is a partial differential equation of the second order in $3N$ variables, and it certainly has solutions for all values of the energy parameter E – real or complex. However, only a selected number of E-values are of physical interest, and they correspond to solutions $\Psi = \Psi(X)$ which satisfy certain physical natural *boundary conditions*. The so-called *closed states* associated with discrete E-values correspond to eigenfunctions $\Psi(X)$ which are absolutely quadratically integrable and which, therefore, belong to the L^2-Hilbert space. On the other hand, the eigenfunctions $\Psi(E, X)$ of the *scattering states* are not absolutely quadratically integrable, and they are instead characterized by the fact that their absolute value $|\Psi(E, X)|$ is *finite* almost everywhere; as a rule, they are associated with *continuous* E-values forming one or more intervals on the real E-axis. These boundary conditions are usually enough to render the spectrum $\{E\}$ of the system under consideration.

Since the L^2 methods are fundamental in solving the eigenvalue problem (55), it is, of course, a disadvantage that the eigenfunctions in the continuum are not directly related to the L^2 space. However, this problem can be removed by using an alternative definition of the boundary conditions for the scattering wave functions Ψ:

$$\Psi \notin L^2, \quad \Psi(E) = \frac{d\Phi(E)}{dE}, \quad \Phi \in L^2, \tag{56}$$

which implies that Ψ is a derivative – in the sense of Lebesgue – of an element Φ which belongs to L^2. This definition implies that one has a limiting procedure:

$$\Psi(E) = \lim_{n\to\infty} \frac{\Phi(E+1/n) - \Phi(E)}{(1/n)} = \lim_{n\to\infty} \Psi_n(E), \tag{57}$$

where $\Psi_n(E) \in L^2$. This result implies that

$$|\Psi(X,E) - \Psi_n(X,E)| < \varepsilon, \quad \text{whenever } n > N(\varepsilon, X), \tag{58}$$

which again implies that one has a non-uniform point-by-point convergence in the absolute value. Hence, the scattering wave function $\Psi(E)$ is a limit point of the L^2 Hilbert space, not in the norm but in absolute value in the sense of (58).

To illustrate the validity of this approach, for a moment we will consider the momentum operator $p = (h/2\pi i)\partial/\partial x$ as an example. The eigenvalue relation

$$p\psi = \lambda\psi \tag{59}$$

has the formal solution $\psi = A\exp[(2\pi i/h)\lambda x]$ for all values of λ which are real or complex, but it is always outside L^2. The absolute value $|\psi|$ will remain finite, if and only if the parameter λ is real. One further has the two relations

$$\Phi(\lambda) = \int_0^\lambda \psi(\lambda)\,d\lambda = A\,\frac{\exp[(2\pi i/h)\lambda x] - 1}{(2\pi i/h)x}; \tag{60}$$

$$|\Phi(\lambda)|^2 = (h/2\pi)^2 |A|^2 \,\frac{2\{1 - \cos(2\pi/h)\lambda x\}}{x^2}, \tag{61}$$

which indicate that $\Phi(\lambda)$ belongs to the L^2-space if and only if $-\infty < \lambda < +\infty$. Hence, the spectrum $\{\lambda\}$ is continuous, filling the entire real axis.

The validity of this approach is essential, since it implies that one may use all the ordinary L^2 methods not only for studying the discrete closed states, but also for treating the scattering states in the continuum. In using the Rayleigh–Ritz method by means of a finite basis of order M, all of the M approximate solutions are in L^2, but some of them correspond to approximate discrete eigenvalues and some of them to approximate continuous eigenvalues. If the number M is increased, the former show a great deal of stability and converge slowly to the exact discrete eigenvalues, whereas the latter change more quickly and tend to give the best approximate scattering functions expressible in terms of the finite basis under consideration. By studying the stability of the approximate eigenvalues, one may even treat discrete eigenvalues which are embedded in the continuum. It is clear that, even if one understands the properties of the approximate eigenvalues rather well from a practical point of view, much more research is needed in clarifying the mathematical convergence properties of the approximate eigenvalues of the Rayleigh–Ritz scheme when M goes to infinity.

However, if one is interested essentially in the low-lying discrete states, all these extra approximate scattering eigenvalues represent an unnecessary complication, and one is usually better off by using resolvent methods or partitioning techniques [4] and by directly aiming at obtaining the discrete eigenvalues desired. Another advantage of this approach is that the characteristic equation is replaced by the *reduced characteristic equation*, which means that all multiple eigenvalues are reduced to single ones. On the other hand, if one is interested in a fixed specific scattering energy E in the continuum, one is better off by using, for example, the Hulthén–Kohn variational principle [5].

In concluding this section, we observe that – from the point of view of the "economy of thinking" – the theory of the abstract Hilbert space provides a marvellous tool since, if one has proven a theorem in this space, it is immediately valid also in all the realizations.

2.7. COMPUTER CONVERGENCE

This paper would be incomplete if we did not say a few words about the practical type of convergence which one meets in computational theoretical chemistry, particularly in connection with iteration procedures. For instance, in the HF-scheme and the MC-SCF method, a computation is said to be *self-consistent* if the numerical result with a specified number of significant figures does not change with further iterations and the iterations have *converged* on the computer. This definition implies that a series of numbers $a_1, a_2, a_3, a_4, \ldots, a_n, \ldots$ becomes practically convergent when, for a specified fixed m, one has

$$|a_{n+m} - a_n| < \varepsilon, \quad \text{whenever} \quad n > N(\varepsilon). \tag{62}$$

This criterion goes over into mathematical convergence if and only if it is valid for *all* values of m. On the other hand, if it is valid only for $m \leq M$, the sequence $\{a_n\}$ may very well be mathematically divergent! A simple example is provided by the divergent series

$\Sigma_n(1/n)$. This result implies that, if a sequence $\{a_n\}$ is computer convergent, one may never find out whether it is mathematically convergent or not by means of a finite number of operations on the computer. In order to solve this problem, one thus has to carry out a careful mathematical analysis of the iteration procedure itself. This means that, at the same time as we in the future are frequently going to utilize the properties of "computer convergence", some research ought to be done as to the question of the true convergence of the iteration procedure. In this connection, it should also be remembered that sometimes strongly divergent procedures are very useful in determining the solutions by considering the "point" from which the procedure diverges [6]. A typical example is provided by the Hartree–Fock method for some negative ions such as, for example, F^- and Cl^-.

3. Conclusions

It is obvious that is does not take too much rigorous mathematics to treat the convergence problems occurring in the theory of the abstract Hilbert space and to apply them to the various realizations: the infinite vector space $\mathcal{H}_0$, the L^2 space, and the Hilbert–Schmidt spaces built on these realizations. By accomplishing this goal, one may raise theoretical chemistry to a higher level of strictness, which is certainly desirable. At the same time, one must remember that any violation of the rules of mathematics may not be used to improve the agreement between the theoretical results and the experimental experience, that improved strictness does not necessarily imply improved results, and that good agreement with the experiments is a necessary but by no means sufficient criterion for the accuracy of the theory. In this paper, we have tried to show that, in theoretical chemistry, there are certainly many different definitions of the concept of convergence which are useful in a variety of mathematical methods and applications.

References

[1] E. Schrödinger, Ann. der Physik 79(1926)361;
W. Heisenberg, Z. Physik 33(1925)879;
M. Born, W. Heisenberg and P. Jordan, Z. Physik 35(1926)557;
P.A.M. Dirac, Proc. Roy. Soc. London A113(1926)621;
E. Schrödinger, Ann. der Physik 79(1926)734.

[2] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer-Verlag, Berlin, 1932).

[3] P.O. Löwdin, Int. J. Quant. Chem. Symp. QC16(1983)485.

[4] For a survey see, for example, P.O. Löwdin, Int. J. Quant. Chem. Symp. 21(1982)69.

[5] L. Hulthén, K. Fysiogr. Sällsk. Lund Förhandl. 14, 21(1844); ibid. 14, 8(1944); C.R. X^{me} Congres des Math. Scan. 1946, Copenhagen (1947), p. 201; Ark. Mat., Fysik o. Astr. 35a, 25(1948);
W. Kohn, Phys. Rev. 74(1948)1763;
S.S. Huang, Phys. Rev. 76(1949)477, 867;
P.O. Löwdin and A. Sjölander, Ark. Mat., Fysik o. Astr. 3, 11(1951).

[6] E. Schröder, Math. Ann. 2(1870)317;
See also P.O. Löwdin, J. Mol. Spec. 10(1963)12, particularly appendix A.