International Journal of Theoretical Physics, Vol. 13, No. 5 (1975), pp. 323-336

A New Differential Calculus on a Complex Banach Space with Application to Variational Problems of Quantum Theory

C. S. SHARMA

Department of Mathematics, Birkbeck College, University of London, England

and

I. REBELO

Mathematical Institute, Oxford University, England

Received: 21 *March* 1975

Abstract

It is proved that a semilinear function on a complex banach space is not differentiable according to the usual definition of differentiability in the calculus on banach spaces. It is shown that this result makes the calculus largely inapplicable to the solution of variational problems of quantum mechanics. A new concept of differentiability called semi-differentiability is defined. This generalizes the standard concept of differentiability in a banach space and the resulting calculus is particularly suitable for optimizing real-valued functions on a complex banach space and is directly applicable to the solution of quantum mechanical variational problems. As an example of such application a rigorous proof of a generalized version of a result due to Sharma (1969) is given. In the course of this work a new concept of prelinearity is defined and some standard results in the calculus on banach spaces are extended and generalized into more powerful ones applicable directly to pre-linear functions and hence yielding the standard results for linear functions as particular cases.

1. *Introduction*

The best solutions of problems both in industry and in science have necessarily an optimal character: from the available solutions one picks out the one which maximizes the good aspects (profitability, efficiency, etc., in industry and exactness, simplicity, etc., in science). Where the exact solution of a physical problem is supposed to be given by the solution of a certain differential equation with appropriate boundary conditions and it is not possible or feasible to solve this equation exactly, approximate solutions can often be obtained by finding a functional on the space of acceptable functions which has a stationary value at the exact solution and then finding the stationary value in a suitable subspace. One can obtain a good estimate of the quality of the approximation if the variational method provides also some means of putting bounds on the errors. Optimization problems also arise in physics because laws of nature invariably seem to have an extremum character: example, the principle of least action.

Differential calculus on banach spaces developed in the 1920s by Fréchet and others is the vehicle of modern theory of optimization. Engineers, control scientists, numerical analysts and a variety of other technologists are finding this calculus an indispensable tool of their respective trades. Theoretical physics now is almost completely quantum physics and though quantum physics is formulated in a hilbert space which is necessarily also a banach space, the differential calculus on banach spaces is rarely used by theoretical physicists. The purpose of this paper is to investigate the reasons for the extraordinary neglect of this calculus by theoretical physicists and to propose a modified calculus on complex hilbert spaces which is likely to be directly useful in the solution of optimization problems of quantum physics.

When the banach space is over the real field and is one dimensional, real functionals on the banach space can be looked upon as a real function of one real variable ($f\colon \mathbb{R} \to \mathbb{R}$). The first and higher derivatives of such functions, if they exist, in traditional elementary calculus are functions $f'\colon \mathbb{R} \to \mathbb{R}$ and, in general, $f^{(n)}\colon \mathbb{R} \to \mathbb{R}$. The first derivative in a certain sense measures the rate of change in the value of f as the argument of the function undergoes a small increment. A stationary point (or critical point) of the function is, by definition, a point $x \in \mathbb{R}$ such that $f'(x) = 0$; the nature of the stationary point (maximum, minimum, or point of inflexion) is determined by examining the higher derivatives. When the banach space is over the real field and of dimension n (finite), a real functional on the space can be looked upon as a function $f\colon \mathbb{R}^n \to \mathbb{R}$, that is, a real function of n real variables. Now from any point $x \in \mathbb{R}^n$ one can move in infinitely many directions and the rate of change in $f(x)$ would, in general, be different in different directions. However, if f is sufficiently smooth, in traditional elementary calculus all these infinitely many rates of change are known if one knows the n partial derivatives or, in other words, if one knows the "grad" (defined symbolically as $\Sigma_{i=1}^{n} \hat{i}_i\,(\partial/\partial x_i)$ where $\hat{i}$ is the unit vector in the direction of the ith coordinate axis in a cartesian frame) of the function at x: the directional derivative (the rate of change) of the function in the direction of any unit vector $\hat{a}$ is simply $\hat{a}\,.\,\text{grad}\, f(x)$. It is intuitively obvious (particularly when $n = 3$ and one looks upon $\mathbb{R}^3$ as a copy of the real physical space) that the rate of change of the function is a local property of the function on the space and does not depend on the choice of a coordinate system. In traditional elementary calculus, *grad* is usually first defined with the help of a particular choice of coordinates and then one proves that *grad* $f(x)$, if it exists, is independent of the choice of the coordinate system. The greatest advantage of the calculus on banach spaces is that it avoids much of this wasteful procedure by using methods which are not only coordinate free but valid also for spaces of any dimension (finite or infinite) and are equally applicable to banach spaces over any field (real, complex, or some other). In particular if the banach space has a positive definite hermitian form (a generalization of the elementary "dot product") defined on it, the definition of *grad* does not depend on partial derivatives or on the choice of a particular coordinate system. One has to become familiar with a few new concepts but the analytic power provided by these concepts makes such a task very much worthwhile.

The optimization of functionals has always been an important method for finding approximate solutions in quantum mechanical problems, but as already stated earlier one does not find many examples of the use of the calculus on banach spaces for the study of this method in quantum theory. The reason for this surprising fact is that a semilinear (which is what Dirac calls "antilinear") functional on a complex hilbert space is not differentiable according to the definition of differentiability in the calculus on banach spaces. Since much of the structure of a complex hilbert space originates from the hermitian product which is semilinear in one of the two variables, most of the functionals which occur in quantum mechanical problems are not differentiable. This difficulty can be overcome either by reducing the problem to one in a real hilbert space (Sharma and Rebelo, 1973a, b) where semilinearity becomes identical with linearity and the difficulties vanish or by a suitable modification of the concept of differentiability so that it may become possible for semilinear functionals to have derivatives and then the resulting modified calculus can be used directly to study optimization problems in complex hilbert spaces. We follow the latter alternative in this paper. By suitably amending the definition of differentiability we obtain a new concept which we call semidifferentiability and which leads to a new calculus on complex banach spaces. We also study some of the applications of this new calculus to optimization problems in quantum theory.

In the next Section we recapitulate some definitions and results which are well known and generalize them to get the new concept of prelinearity and results applicable to prelinear functions. Where the proof of the generalized version is very similar to that of the standard version of a proposition we merely give the crucial step. In Section 3 we prove that a semilinear functional on a complex banach space, except in the trivial case, is not differentiable and then we modify the definition of differentiability and obtain some basic results of the new calculus. In Section 4 we give some applications of this calculus.

2. *Recapitulation and Generalization of Some Standard Definitions and Results*

We collect below the definitions [D], propositions [P] and observations [O] we need for our work. The concept of prelinearity is new and therefore all propositions involving this concept are generalizations of established results. In what follows, all vector spaces unless specifically stated otherwise are over the complex field.

[D1] A function f from a vector space V_1 to another vector space V_2 is said to be *linear* if

$$f(x+y)=f(x)+f(y) \qquad \forall x,y \in V_1$$

and

$$f(\alpha x) = \alpha f(x) \qquad \forall \alpha \in \mathbb{C} \text{ and } \forall x \in V_1$$

[D2] A function f from a vector space V_1 to another vector space V_2 is said to be *semilinear* if

$$f(x + y) = f(x) + f(y) \qquad \forall x, y \in V_1$$

and

$$f(\alpha x) = \bar{\alpha} f(x) \qquad \forall \alpha \in \mathbb{C} \text{ and } \forall x \in V_1$$

[D3] A function f from a vector space V_1 to a normed space V_2 is said to be *prelinear* if

$$f(x + y) = f(x) + f(y)$$

and

$$\|f(\alpha x)\| = |\alpha| \|f(x)\| \qquad \forall \alpha \in \mathbb{C} \text{ and } \forall x \in V_1$$

[O1] Whenever the range of a linear or a semilinear function lies in a normed space, the function satisfies the criterion of being prelinear.

[D4] A *functional* f on a vector space V is a function from V to $\mathbb{C}$.

[D5] A linear (or semilinear or prelinear) function f from a normed space V_1 to another normed space V_2 is said to be *bounded* if there exists a real number M such that

$$\|f(x)\| \leqslant M \|x\| \qquad \forall x \in V_1$$

[D6] A linear (or semilinear or prelinear) function f from a normed space V_1 to another normed space V_2 is said to be *continuous at a point* $x_0 \in V_1$ if given a real positive number $\mathscr{E}$ there exists a real positive number δ such that

$$\|x - x_0\| < \delta \Rightarrow \|f(x) - f(x_0)\| < \mathscr{E}$$

[P1] Let f be a prelinear function from a normed space N_1 to another normed space N_2. The following assertions about f are equivalent:

(a) f is continuous at a point $x_0 \in N_1$
(b) f is bounded
(c) f is continuous at every point $x \in N_1$

Proof. Suppose f is continuous at x_0. This implies that given $\epsilon > 0$, there exists a real $\delta > 0$ such that $\|x - x_0\| < \delta \Rightarrow \|f(x - x_0)\| < \epsilon$. But $\|x_0 - (x_0 - (\delta/2\|x\|)x)\| = \delta/2 < \delta \Rightarrow \|f(x_0 - x_0 + (\delta/2\|x\|)x)\| < \epsilon \Rightarrow \|f(x)\| < 2\epsilon/\delta \; \|x\|$. Hence f is bounded by $2\epsilon/\delta$. Thus $(a) \Rightarrow (b)$.

Suppose f is bounded. Hence there exists a real number M such that $\|f(x - x_0)\| \leqslant M \|x - x_0\|$. Given $\epsilon > 0$ take $\delta = \epsilon/M$. Then $\|x - x_0\| < \delta \Rightarrow \|f(x - x_0)\| < \epsilon$. Hence f is continuous at x where x is any point of N_1. Thus (b) $\Rightarrow$ (c).

(c) obviously implies (a).

[D7] A normed vector space which is complete in the metric topology induced by the norm is called a *banach space.*

[D8] Let N_1 and N_2 be two normed spaces. The space of all functions from N_1 to N_2 will be denoted by (N_1, N_2). The subspace of all bounded prelinear, linear and semilinear functions from N_1 to N_2 will be denoted by $\mathscr{PL}(N_1, N_2)$, $\mathscr{L}(N_1, N_2)$ and $\mathscr{SL}(N_1, N_2)$ respectively.

[P2] Let N_1 and N_2 be normed spaces. Then $\mathscr{PL}(N_1, N_2)$ is a normed space. If in addition N_2 is a banach space, then so also is $\mathscr{PL}(N_1, N_2)$.

Proof. The verification of the fact that $\mathscr{PL}(N_1, N_2)$ is a vector space is trivial. Let $f \in \mathscr{PL}(N_1, N_2)$. It is easy to verify that

$$\|f\| = \sup_{\substack{\|x\|=1 \\ x \in N_1}} \|f(x)\|$$

defines a norm on $\mathscr{PL}(N_1, N_2)$, hence $\mathscr{PL}(N_1, N_2)$ is a normed space. Finally suppose that N_2 is a banach space. Let $\{f_n\}$ be a cauchy sequence in $\mathscr{PL}(N_1, N_2)$. For any fixed $x \in N_1$ $\|f_n(x) - f_m(x)\| \leqslant \|f_n - f_m\| \, \|x\| \Rightarrow \{f_n(x)\}$ is a cauchy sequence in N_2 and hence converges to some point $y \in N_2$. Define a function f from N_1 to N_2 by $f(n) = \lim f_n(x)$. It is then easy to verify that $f \in \mathscr{PL}(N_1, N_2)$ and f is the limit of $\{f_n\}$.

[O2] Since both linear and semilinear maps are also prelinear [P1] and [P2] are automatically valid for such maps.

[D9] A *sesquilinear form* s on a complex vector space V is a function s: $V \times V \to \mathbb{C}$ which is a linear functional on V for any fixed first member of an element of $V \times V$ and which is a semilinear functional on V for any fixed second member of an element of $V \times V$ (In pure mathematics texts the convention used is that a sesquilinear form is linear in its first argument and semilinear in its second and is exactly opposite to that used in the preceding definition which however conforms with the convention used almost universally in quantum theory texts.)

[D10] A sesquilinear form s on a vector space V is said to be *hermitian* if $s(x, y) = \overline{s(y, x)} \quad \forall (x, y) \in V \times V$.

A hermitian form s on a vector space V is said to be *positive* if $h(x, x) \geqslant 0$ $\forall x \in V$. A positive hermitian form h on a vector space V is called positive definite if $h(x, x) = 0 \Leftrightarrow x = 0$.

[P3] Let $\langle \, , \rangle$ be a positive definite hermitian form on a complex vector space V. Then

$$|\langle x, y\rangle|^2 \leqslant \langle x, x\rangle\langle y, y\rangle \qquad \forall x, y \in V \text{ (cauchy-schwarz inequality)}$$

Proof. This is an elementary consequence of the assumed positiveness of the hermitian form. If either $x = 0$ or $y = 0$, there is nothing to prove. Suppose $y \neq 0$, expand $0 \leqslant \langle x + ty, x + ty\rangle$ and take $t = -\langle y, x\rangle / \langle y, y\rangle$. The inequality follows.

[P4] Let $\langle \, , \rangle$ be a positive definite hermitian form on a complex vector space V. Then $x \mapsto \langle x, x\rangle^{1/2}$ defines a norm on V.

Proof. An elementary consequence of [D10] and [P3].

[D11] A vector space over $\mathbb{C}$ with a positive definite hermitian form defined on it is called a *prehilbert space.*

[D12] A prehilbert space which is complete in the metric topology induced by its hermitian form is called a *hilbert space* (the hermitian form induces a norm and the norm induces a metric).

[P5] A hilbert space is necessarily a banach space.

Proof. Follows immediately from [D7] and [D12].

[P6] (The Riesz representation theorem). Let $\mathscr{H}$ be a hilbert space. Let $\Phi_y(x) = \langle y, x\rangle$. Clearly $\Phi_y \in \mathscr{L}(\mathscr{H}, \mathbb{C})$. The correspondence $y \mapsto \Phi_y$ is a norm preserving semilinear isomorphism between $\mathscr{H}$ and $\mathscr{L}(\mathscr{H}, \mathbb{C})$. ($\mathscr{L}(\mathscr{H}, \mathbb{C})$ is also called the dual of $\mathscr{H}$).

Proof. $\|\Phi_y\| = \sup\limits_{\|x\|=1} |\langle y, x\rangle| \leqslant \|y\|$

Also $\|\Phi_y\| = \sup\limits_{\|x\|=1} |\langle y, x\rangle| \geqslant |\langle y, y/\|y\|\rangle| = \|y\|$

Hence the correspondence is norm preserving. Semilinearity is trivial to verify. To prove that the correspondence is onto, we must show that if $\Phi \in \mathscr{L}(\mathscr{H}, \mathbb{C})$, there exists an $y \in \mathscr{H}$ such that $\Phi(x) = \langle y, x\rangle$. If $Ker\ \Phi = \{x \in \mathscr{H}\colon \Phi(x) = 0\}$ is the whole of $\mathscr{H}$, then $y = 0$ satisfies the requirement. If $Ker\ \Phi \neq \mathscr{H}$, let $Ker\ \Phi = \mathscr{G}$. It is evident from the definition that $\mathscr{G}$ is a linear manifold in $\mathscr{H}$ and it follows from the continuity of Φ that it is closed. Hence it is a subspace of $\mathscr{H}$. Since $\mathscr{G} \neq \mathscr{H}, \mathscr{G}^\perp$ contains a nonzero vector z. It is easy to verify that $y = (\overline{\Phi(z)}/\|z\|^2)z$ satisfies the requirement. To prove that the correspondence is one-one, suppose it is not, then

$$\Phi_y(x) = \langle y, x\rangle = \langle y', x\rangle = \Phi_{y'}(x) \qquad \forall x \in \mathscr{H}$$

$$\Rightarrow \langle y - y', x\rangle = 0 \qquad \forall x \in \mathscr{H}$$

Take $x = y - y'$ and conclude that $y = y'$.

[P7] Let $\mathscr{H}$ be a hilbert space. Let $\Phi_y(x) = \langle x, y\rangle$. Clearly $\Phi_y \in \mathscr{SL}(\mathscr{H}, \mathbb{C})$. The correspondence $y \to \Phi_y$ is a norm preserving linear isomorphism between $\mathscr{H}$ and $\mathscr{SL}(\mathscr{H}, \mathbb{C})$.

Proof. This is an immediate corollary of [P6].

[D13] Let $\mathscr{B}_1$ and $\mathscr{B}_2$ be two banach spaces. Let $f\colon \mathscr{B}_1 \to \mathscr{B}_2$. f is said to be *differentiable at a point* $x \in \mathscr{B}_1$ if there exists an $f'_x \in \mathscr{L}(\mathscr{B}_1, \mathscr{B}_2)$ such that

$$\lim_{\|u\| \to 0} \frac{\|f(x+u) - f(x) - f'_x(u)\|}{\|u\|} = 0.$$

If f is differentiable at every point of $\mathscr{B}_1$, then it is said to be *differentiable in* $\mathscr{B}_1$. In such a case f' which assigns to each $x \in \mathscr{B}_1$ the derivative f'_x at that point is called the *derivative* of f. The derivative of f', if it exists, is denoted by f'' and is called the *second derivative.* Higher derivatives are defined analogously.

[O3] We have defined differentiability for a function defined on the whole of $\mathscr{B}_1$. It is, however, customary to define this for functions whose domains are

open subsets of $\mathscr{B}_1$. The whole of $\mathscr{B}_1$ is, of course, an open subset of $\mathscr{B}_1$.
[O4] According to Cartan's (1971) definition f is differentiable at x if f is continuous at x and there exists a linear map from $\mathscr{B}_1$ to $\mathscr{B}_2$ satisfying the requirement in [D13]. Our definition requires f'_x to be continuous (cf. [P1]). Since the continuity of f implies that of f'_x and vice versa the two definitions are equivalent. For a finite dimensional $\mathscr{B}_1$ linear maps from $\mathscr{B}_1$ to $\mathscr{B}_2$ are necessarily continuous, hence the assumption of the continuity of f becomes superfluous in this case. Our definition in all cases requires the differentiability to imply continuity.
[O5] Let $f: \mathbb{R}^n \to \mathbb{R}$. In traditional elementary calculus of many variables only directional (or partial) derivatives are defined and the directional derivative f'_a in the direction of a vector a is a function from $\mathbb{R}^n$ to $\mathbb{R}$ and $f'_a(x)$ is the value f'_a takes at x and is a number. In our definition a derivative f' is defined and is a function from $\mathbb{R}^n$ to $\mathscr{L}(\mathbb{R}^n, \mathbb{R})$, f'_x is the value f' takes at x and is a linear map from $\mathbb{R}^n$ to $\mathbb{R}$ and $f'_x(y)$ $(y \in \mathbb{R}^n)$ is a number. The value $f'_a(x)$ of the directional derivative of traditional calculus in terms of our derivative is simply $f'_n(a/\|a\|)$. Similar remarks apply about higher derivatives.
[O6] It follows from the definition of the derivative and [P6] that if a functional f on a hilbert space is differentiable at a point $x \in \mathscr{H}$ then since f'_x is a linear functional there is a unique element $y_x \in \mathscr{H}$ such that $f'_x(u) = \langle y'_x, u \rangle$.
[D14] From [O6] we know that if a functional on a hilbert space is differentiable then the derivative f'_x at x corresponds to a unique vector y_x in such a way that $f'_x(u) = \langle y_x, u \rangle$. This unique vector y_x is called the *gradient* of f at x and is denoted by grad $f(x)$.
[O7] The Riesz representation theorem gives us a coordinate free definition of the gradient for functionals defined on a hilbert space. In this approach it is no longer necessary to prove the invariance of grad with respect to a change in coordinates.

The theory developed so far is equally applicable to both real and complex hilbert spaces. Nevertheless, for applications to quantum mechanics it needs important modifications. To formulate the mathematics involved in the applications, it is necessary to have some knowledge of the theory of measures and integrals which, in what follows, will be assumed (definitions of terms which are used without explicit definition will be found in Halmos (1950, 1957)). It will no longer be possible to indicate even briefly the main ideas involved in the proofs of the established results we need as many of these proofs are rather involved and complicated. However, most physicists are aware of the spectral theorem which is the main result we need and it is, therefore, hoped that what follows would be intelligible to the reader even though some of the definitions and proofs have been omitted.
[D15] Let $\boldsymbol{B}$ be the borel algebra of the reals. Let $\mathscr{E}$ be the set of projections on the subspaces of a hilbert space. A *spectral measure* is a function $E: \boldsymbol{B} \to \mathscr{E}$ $(\lambda \mapsto E_\lambda)$ such that $E_{\mathbb{R}} = 1, E_\phi = 0$ and for a disjoint sequence of sets $\{\lambda_i\}$ in $\boldsymbol{B}$

$$E_{\bigcup_i \lambda_i} = \sum_i E_{\lambda_i}$$

[D16] A linear continuous function from a hilbert space $\mathscr{H}$ to itself is called an *endomorphism* on $\mathscr{H}$.

[D17] Let A $(x \mapsto Ax)$ be an endomorphism on a hilbert space $\mathscr{H}$. A is said to be *self-adjoint* if $\langle x, Ay\rangle = \langle Ax, y\rangle \,\forall x, y \in \mathscr{H}$.

[D18] Let A be an endomorphism on $\mathscr{H}$. The *spectrum* $\Lambda(A)$ of A is a subset of $\mathbb{C}$ defined by

$$\Lambda(A) = \{\lambda \in \mathbb{C} : A - \lambda I \text{ is not invertible}\}.$$

[P8] The spectrum of a self-adjoint endomorphism on $\mathscr{H}$ is a subset of $\mathbb{R}$. Furthermore corresponding to each self-adjoint endomorphism on $\mathscr{H}$ there is a unique spectral measure E on the borel algebra $\boldsymbol{B}$ of the reals such that

$$\lambda \in \boldsymbol{B} \qquad \text{and} \qquad \lambda \cap \Lambda(A) = \phi \Rightarrow E_\lambda = 0$$

and $\int_\Lambda E_\lambda = A$.

Proof. See Halmos (1957).

3. *A Modified Calculus on a Complex Hilbert Space*

We first prove a proposition which asserts that except in the trivial case a semilinear function from a complex banach space to another normed space is not differentiable. This result explains why the calculus on banach spaces cannot be used directly for the study of optimization in quantum mechanics.

[P9] Let $\mathscr{B}$ be a complex banach space and let N be any normed space. Let $f \in \mathscr{SL}(\mathscr{B}, N)$, then f is differentiable if and only if $f = 0$ which is the trivial case.

Proof. Suppose that f is differentiable at x and f'_x is the derivative at x. Let u be any nonzero element of $\mathscr{B}$, then by the definition of differentiability

$$\lim_{\substack{t \in \mathbb{R} \\ t \to 0}} \frac{\|f(x + tu) - f(x) - f'_x(tu)\|}{\|tu\|} = 0$$

$$\Rightarrow \frac{\|f(u) - f'_x(u)\|}{\|u\|} = 0$$

$$\Rightarrow \|f(u) - f'_x(u)\| = 0 \qquad \text{since} \quad \|u\| \neq 0$$

$$\Rightarrow f(u) = f'_x(u) \tag{3.1}$$

Also from the definition of differentiability

$$\lim_{\substack{t \in \mathbb{R} \\ t \to 0}} \frac{\|f(x + itu) - f(x) - f'_x(itu)\|}{\|itu\|} = 0$$

$$\Rightarrow f(u) = -f'_x(u) \tag{3.2}$$

Equations (3.1) and (3.2) together imply that $f(u) = 0$, but u is any nonzero element of $\mathscr{B}$. Hence $f = 0$.

If $f = 0$, it is, of course, differentiable with $f' = 0$. We are finished.

All the functionals on a hilbert space which are defined with the help of the hermitian product are semilinear in one of the variables and so the great majority of functionals occurring in quantum theory are partly semilinear and therefore not differentiable. However, semilinear functionals are related to the linear ones by complex conjugation and both are equally smooth. By a modification of the usual definition of differentiability it is possible to develop a differential calculus which can deal with functionals which are semilinear. To avoid confusion we have decided to call the new differentiability by the name of semidifferentiability partly because differentiability implies semidifferentiability and partly because it is closely related to semilinear functions.

[D19] Let $\mathscr{B}_1$ and $\mathscr{B}_2$ be two complex banach spaces. A function f: $\mathscr{B}_1 \to \mathscr{B}_2$ is said to be *semidifferentiable* at a point $x \in \mathscr{B}_1$, if there exists a function $f_x^{(s)} \in \mathscr{L}(\mathscr{B}_1, \mathscr{B}_2) \oplus \mathscr{SL}(\mathscr{B}_1, \mathscr{B}_2)$ such that

$$\lim_{\|n\| \to 0} \frac{\|f(x+u) - f(x) - f_x^{(s)}(u)\|}{\|u\|} = 0$$

The function $f_x^{(s)}$, if it exists, is called the *semiderivative* of f at x. If the function f is semidifferentiable at each point of $\mathscr{B}_1$, it is said to be semidifferentiable in $\mathscr{B}_1$ and the rule which assigns to each point x the semiderivative of f at that point is called the semiderivative of f in $\mathscr{B}_1$ and is denoted by $f^{(s)}$. Higher derivatives are defined analogously. For example, the second derivative

$$f^{(2s)} \in (\mathscr{B}_1, \mathscr{L}(\mathscr{B}_1, \mathscr{L}(\mathscr{B}_1, \mathscr{B}_2) \oplus \mathscr{SL}(\mathscr{B}_1, \mathscr{B}_2)) \oplus \mathscr{SL}(\mathscr{B}_1, \mathscr{L}(\mathscr{B}_1, \mathscr{B}_2) \oplus \mathscr{SL}(\mathscr{B}_1, \mathscr{B}_2))).$$

[O8] We will be mostly concerned with functionals on hilbert spaces, that is, the special case where $\mathscr{B}_1 = \mathscr{H}$ and $\mathscr{B}_2 = \mathbb{C}$. For this case with the help of [P6] and [P7] we immediately see that if $f_x^{(s)}$ exists, then it has a unique representation of the form

$$f_x^{(s)}(u) = \langle y_x, u \rangle + \langle u, z_x \rangle$$

[D20] The unique pair of vectors y_x and z_x of [O8] are called *lingrad* and *semilingrad* of f at x.

[D21] A *stationary point* of a semidifferentiable functional f on $\mathscr{H}$ is defined to be a point $x \in \mathscr{H}$ such that

$$f_x^{(s)} = 0$$

[O9] $f_x^{(s)} = 0 \Rightarrow \text{lingrad} f(x) = \text{semilingrad} f(x) = 0$

[P10] Let A be an endomorphism on $\mathscr{H}$ and let $a \in \mathscr{H}$. Let f_1, f_2 and f_3 be functionals on $\mathscr{H}$ defined by

$$f_1(x) = \langle x, Aa \rangle$$
$$f_2(x) = \langle a, Ax \rangle$$
$$f_3(x) = \langle x, Ax \rangle$$

then

$$\operatorname{lingrad} f_1(x) = 0; \ \operatorname{semilingrad} f_1(x) = Aa$$
$$\operatorname{lingrad} f_2(x) = A^*a; \ \operatorname{semilingrad} f_2(x) = 0$$
$$\operatorname{lingrad} f_3(x) = A^*x; \ \operatorname{semilingrad} f_3(x) = Ax$$

Proof. The verifications, with the help of [D19], are quite straightforward and will be carried through for f_3 only. Set

$$f_{3x}^{(s)}(u) = \langle A^*x, u\rangle + \langle u, Ax\rangle$$

$$\lim_{\|u\|\to 0} \frac{|f_3(x+u) - f_3(x) - f_{3x}^{(s)}(u)|}{\|u\|}$$
$$= \lim_{\|u\|\to 0} \frac{|\langle x+u, A(x+u)\rangle - \langle x, Ax\rangle - \langle x, Au\rangle - \langle A^*x, u\rangle|}{\|u\|}$$
$$= \lim_{\|u\|\to 0} \frac{|\langle u, Au\rangle|}{\|u\|} \leqslant \lim_{\|u\|\to 0} \frac{\|A\|\,\|u\|^2}{\|u\|} = \lim_{\|u\|\to 0} \|A\|\,\|u\| = 0$$

[P11] The second semiderivatives of the three functionals defined in [P10] are given by

$$f_{1x}^{(2s)}(u)(v) = f_{2x}^{(2s)}(u)(v) = 0$$
$$f_{3x}^{(2s)}(u)(v) = \langle u, Av\rangle + \langle v, Au\rangle$$

Proof. Elementary.

[P12] Let $f_1: \mathscr{H} \to \mathbb{C}$ and $f_2: \mathscr{H} \to \mathbb{C}$ be two functionals semidifferentiable at $x \in \mathscr{H}$. Then $F: \mathscr{H} \to \mathbb{C}$ defined by $F(x) = f_1(x) f_2(x)$ is semidifferentiable at x and

$$F_x^{(s)}(u) = f_{1x}^{(s)}(u) f_2(x) + f_1(x) f_{2x}^{(s)}(u)$$

Proof. $$\lim_{\|u\|\to 0} \frac{|F(x+u) - F(x) - f_{1x}^{(s)}(u) f_2(x) - f_1(x) f_{2x}^{(s)}(u)|}{\|u\|}$$
$$= \lim_{\|u\|\to 0} \frac{1}{\|u\|} \,|\, [f_1(x+u)[f_2(x+u) - f_2(x) - f_{2x}^{(s)}(u)]$$
$$+ f_2(x)[f_1(x+u) - f_1(x) - f_{1x}^{(s)}(u)] + f_{2x}^{(s)}(u)[f_1(x+u) - f_1(x)]]\,|$$
$$\leqslant \lim_{\|u\|\to 0} \Bigg[|f_1(x+u)| \frac{|f_2(x+u) - f_2(x) - f_{2x}^{(s)}(u)|}{\|u\|}$$
$$+ |f_2(x)| \frac{|f_1(x+u) - f_1(x) - f_{1x}^{(s)}(u)|}{\|u\|}$$
$$+ f_{2x}^{(s)}\left(\frac{u}{\|u\|}\right) |f_1(x+u) - f_1(x)| \Bigg] = 0$$

where we have used the definition of semidifferentiability and the continuity and boundedness of f_1, f_2 and $f_{2x}^{(s)}$.

[P13] Let $\mathscr{B}_1, \mathscr{B}_2$ and $\mathscr{B}_3$ be three complex banach spaces. Let $f_1: \mathscr{B}_1 \to \mathscr{B}_2$ be semidifferentiable at $x \in \mathscr{B}_1$ and $f_2: \mathscr{B}_2 \to \mathscr{B}_3$ be semidifferentiable at $f_1(x) \in \mathscr{B}_2$. Then $F = f_2 \circ f_1: \mathscr{B}_1 \to \mathscr{B}_3$ is semidifferentiable at x and

$$F_x^{(s)}(u) = f_{2f_1(x)}^{(s)} \circ f_{1x}^{(s)}(u)$$

Proof. A computation very similar to the one used in [P12].

[P14] Let $f: \mathscr{B}_1 \to \mathscr{B}_2$ be a $(n+1)$–times semidifferentiable function and let $\|f^{((n+1)s)}\| \leqslant M$

then

$$\|f(a+u) - f(a) - f_a^{(s)}(u) - \dots - \frac{1}{n!} f_a^{(ns)}(u, \dots^{n \text{ times}}, u)\| \leqslant \frac{M\|u\|^{n+1}}{n!}$$

Proof. A check of Cartan's (1971) proof of the similar formula for a $(n+1)$-times differentiable function shows that semidifferentiability is enough for the validity of the proof.

4. *Applications to Quantum Mechanical Problems*

As an example of the applications of the calculus developed in the preceding Section to quantum mechanical problems, we give a simple, elegant and rigorous proof of a generalization of an earlier result of Sharma (1969) concerning upper and lower bounds for a certain class of quantum mechanical sums.

Let H_0 be a self-adjoint endomorphism on $\mathscr{H}$ and let H_1 be any endomorphism on $\mathscr{H}$. Let $\epsilon_i^{(0)}$ be an eigenvalue of H_0 of finite multiplicity and such that $\Lambda'(H_0) = \{\lambda : \lambda \in \Lambda(H_0) \;\&\; \lambda < \epsilon_i^{(0)}\}$ is a finite set consisting of eigenvalues of finite multiplicities. The labeling index i on the eigenvalues arranges them in a monotonically increasing sequence. Our example concerns the evaluation of a sum of the kind

$$S_{i,r,s} = - \int_{\Lambda(H_0)\setminus\{\epsilon_i^{(0)}\}} (\lambda - \epsilon_i^{(0)} + \alpha)^r \{(\lambda - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^s \|E_\lambda H_1 \Phi_i^{(0)}\|^2 \tag{4.1}$$

where E is the spectral measure of H_0, r and s are integers (positive, negative or zero), α is a real number such that $\epsilon_i^{(0)} - \alpha \notin \Lambda(H_0)$, β is any real number, $\Phi_i^{(0)}$ is a given eigenvector of H_0 belonging to the eigenvalue $\epsilon_i^{(0)}$. As in the earlier work of Sharma (1969) we define positive numbers r′ and r″ by

$$\begin{aligned} r' &= \tfrac{1}{2}(|r| + r) \\ r'' &= \tfrac{1}{2}(|r| - r) \end{aligned} \tag{4.2}$$

Positive numbers s′ and s″ are defined analogously.

With the help of spectral analysis it is not difficult to verify that

$$\begin{aligned} -S_{i,r,s} &= \langle \Psi, (H_0 - \epsilon_i^{(0)} + \alpha)^{|r|}((H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2)^{|s|}(1 - E_{\epsilon_i(0)})\Psi\rangle \\ &= -\langle \Psi, (1 - E_{\epsilon_i(0)})(H_0 - \epsilon_i^{(0)} + \alpha)^{r'}((H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2)^{s'} H_1 \Phi_i^{(0)}\rangle \end{aligned} \tag{4.3}$$

where Ψ is any solution of

$$\begin{aligned} &(H_0 - \epsilon_i^{(0)} + \alpha)^{|r|}((H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2)^{|s|}(1 - E_{\epsilon_i(0)})\Psi \\ &\quad + (1 - E_{\epsilon_i(0)})(H_0 - \epsilon_i^{(0)} + \alpha)^{r'}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s'} H_1 \Phi_i^{(0)} = 0 \end{aligned} \tag{4.4}$$

In fact with our earlier assumption about $\alpha(\epsilon_i^{(0)} - \alpha \notin \Lambda(H_0))$ it is easy to prove that this equation cannot have more than one solution in $\mathscr{H} \ominus \mathscr{H}_{\epsilon_i(0)}$ where $\mathscr{H}_{\epsilon_i(0)}$ is the eigenspace of H_0 belonging to the eigenvalue $\epsilon_i^{(0)}$. This enables one to construct the generalized hylleraas functional

$$\begin{aligned} E_{i,r,s}(\Psi) &= \langle \Psi, (H_0 - \epsilon_i^{(0)} + \alpha)^{|r|}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{|s|}(1 - E_{\epsilon_i(0)})\Psi\rangle \\ &\quad + \langle \Psi, (1 - E_{\epsilon_i(0)})(H_0 - \epsilon_i^{(0)} + \alpha)^{r'}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s'} H_1 \Phi_i^{(0)}\rangle \\ &\quad + \langle \Phi_i^{(0)}, H_1^*(1 - E_{\epsilon_i(0)})(H_0 - \epsilon_i^{(0)} + \alpha)^{r'}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s'}\Psi\rangle \end{aligned} \tag{4.5}$$

This is a symmetrical functional and its lingrad and semilingrad are identical and equation (4.4) is satisfied at its stationary points. However, from [P14] $E_{i,r,s}$ is a minimum if $(H_0 - \epsilon_i^{(0)} + \alpha)^{|r|}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{|s|}$ is a positive operator and maximum if this operator is negative. If $|r|$ is even then $E_{i,r,s}$ is a minimum and provides an upper bound for the sum. However, in general, $E_{i,r,s}$ is neither an upper bound nor a lower bound. Suppose there exists a positive integer N such that for any integer $t \geqslant N$

$$\{\epsilon_t^{(0)} - \epsilon_i^{(0)} + \alpha\}^{|r|} > 0. \tag{4.6}$$

In order to obtain an upper bound we now define a modified hylleraas functional by

$$E'_{i,r,s} = E_{i,r,s} - \sum_{\substack{n<N \\ n\neq i}} (\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^{|r|}\{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{|s|} \times$$

$$\left\| E_{\epsilon_n(0)}\Psi + \frac{E_{\epsilon_n(0)} H_1 \Phi_i^{(0)}}{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^{r''}\{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s''}} \right\|^2 \tag{4.7}$$

By considering the second semiderivative of the functional we see that the stationary point of this functional is a minimum if

$$(H_0 - \epsilon_i^{(0)} + \alpha)^{|r|}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{|s|} - \sum_{\substack{n<N \\ n=i}} (\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^{|r|} \{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{|s|} E_{\epsilon_n(0)} \quad (4.8)$$

is a positive operator, but this is, of course, so because of the way $E'_{i,r,s}$ has been defined. At the extremum of $E'_{i,r,s}$ we have

$$\begin{aligned}(H_0 - \epsilon_i^{(0)} + \alpha)^{|r|}&\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{|s|}(1 - E_i^{(0)})\Psi \\ &+ (1 - E_{\epsilon_i^{(0)}})(H_0 - \epsilon_i^{(0)} + \alpha)^{\gamma'}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s'} H_1 \Phi_i^{(0)} \\ &- \sum_{\substack{n<N \\ n\neq i}} (\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^{|r|}\{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{|s|} \times \\ &\left(E_{\epsilon_n(0)}\Psi + \frac{E_{\epsilon_n(0)} H_1 \Phi_i^{(0)}}{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^{r''}\{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s''}} \right) = 0 \quad (4.9)\end{aligned}$$

It is not difficult to see that solutions of equations (4.4) and (4.9) have unique and identical components in $\mathscr{H} \ominus (\mathscr{H}_{\epsilon_i^{(0)}} \bigoplus_{\substack{n<N \\ n\neq c}} \mathscr{H}_{\epsilon_n^{(0)}})$ and their components in $\mathscr{H}_{\epsilon_i(0)}$ are completely arbitrary. Further a solution of equation (4.9) can have an arbitrary component in $\bigoplus_{\substack{n<N \\ n\neq i}} \mathscr{H}_{\epsilon_n(0)}$ also. The values which the two functionals $E_{i,r,s}$ and $E'_{i,r,s}$ take at any pair of solutions of equations (4.4) and (4.9) are identically same with the value of the sum we wish to evaluate.

In order to obtain a lower bound, we now define yet another modified hylleraas functional by

$$\begin{aligned}E''_{i,r,s} = E_{i,r,s} &- (\epsilon_N^{(0)} - \epsilon_i^{(0)} + \alpha)^{-r''}\{(\epsilon_N^{(0)} - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{-s''} \times \\ &[\langle \Psi, (H_0 - \epsilon_i^{(0)} + \alpha)^{r' + 2r''}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s' + 2s''}(1 - E_{\epsilon_i(0)})\Psi\rangle \\ &+ \langle (1 - E_{\epsilon_i(0)}) H_1 \Phi_i^{(0)}, (H_0 - \epsilon_i^{(0)} + \alpha)^{r}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s'} H_1 \Phi_i^{(0)}\rangle \\ &+ \langle (H_0 - \epsilon_i^{(0)} + \alpha)^{r'' + r'}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s'' + s'}\Psi, (1 - E_{\epsilon_i(0)}) H_1 \Phi_i^{(0)}\rangle \\ &+ \langle (1 - E_{\epsilon_i(0)}) H_1 \Phi_i^{(0)}, (H_0 - \epsilon_i^{(0)} + \alpha)^{r' + r''}\{(H_0 - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s'' + s'}\Psi\rangle\} \\ &+ \sum_{\substack{n<N \\ n\neq i}} \left[\frac{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^{r' + 2r''}\{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s' + s''}}{(\epsilon_N^{(0)} - \epsilon_i^{(0)} + \alpha)^{r''}\{(\epsilon_N^{(0)} - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s''}} \right. \\ &\left. - (\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^{|r|}\{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{|s|} \right] \times \\ &\left\| E_{\epsilon_n(0)}\Psi + \frac{E_{\epsilon_i(0)} H_1 \Phi_i^{(0)}}{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^{r''}\{(\epsilon_n^{(0)} - \epsilon_i^{(0)} + \alpha)^2 + \beta^2\}^{s''}} \right\|^2 \quad (4.10)\end{aligned}$$

This functional has been constructed in such a way that the component of any of its stationary points in $\mathscr{H} \ominus (\mathscr{H}_{\epsilon_i(0)} \bigoplus_{\substack{n<N \\ n\neq i}} \mathscr{H}_{\epsilon_n(0)})$ is identical with the corresponding component of the solutions of equations (4.4) and (4.9) and the value which the functional takes at any stationary point is the value of the sum we wish to evaluate and this value is a maximum. These claims can be easily verified by considering the first and second semiderivatives of the functional.

It should, of course, be pointed out that the positivedefiniteness (negative-definiteness) of the second derivative at a stationary point does not guarantee that the functional is a minimum (maximum) in an infinite dimensional vector space (a counterexample demonstrating the truth of this assertion has been worked out by S. Mare who will publish the result in due course and will report it also in his thesis which he plans to submit to the University of London for a Ph.D.). However, in the functionals considered above third and higher derivatives vanish identically and in such cases the positivedefiniteness (negative-definiteness) of the second derivative (semiderivative) is enough to guarantee that the functional is a minimum (maximum).

The proof given here is more rigorous and the results are more general than those of Sharma (1969). The generalization removes the requirement of nondegeneracy of the eigenvalues. Furthermore, the method is completely coordinate free. However, it must be pointed out that it has been assumed that H_0 is a self-adjoint endomorphism whereas H_0 in quantum mechanical problems is usually an unbounded operator which cannot be continuous at any point. However, if P is the projection on any finite dimensional subspace then PH_0P is an endomorphism and since a variational calculation, because of its very nature, has to be carried out in a finite dimensional subspace, this is all we need.

Acknowledgments

One of us (I.R.) would like to express her gratitude to the International Federation of University Women for the Dorothy Leet Grant Award and to the British Federation of University Women for their Memorial Award.

References

Cartan, H. (1971). *Differential Calculus.* Hermann, Paris.
Halmos, P. (1950). *Measure Theory.* Van Nostrand, New York.
Halmos, P. (1957). *Introduction to Hilbert Space and the Theory of Spectral Multiplicity.* Chelsea, New York.
Sharma, C. S. (1969). *J. Phys. A.* (London), **2**, 413.
Sharma, C. S., and Rebelo, I. (1973a). *J. Phys. A.* (London), **6**, 459.
Sharma, C. S., and Rebelo, I. (1973b). *Phys. Letters* (Amsterdam), **44A**, 29.