International Journal of Theoretical Physics, Vol. 22, No. 2, 1983

Calculus on Complex Banach Spaces

J. Pian and C. S. Sharma

Department of Mathematics, Birkbeck College, Malet Street, London WC1E 7HX.

Received April 6, 1981; revised June 15, 1981

By using the techniques of modern functional analysis, a variety of new concepts have been developed and new results proved which extend considerably the new calculus on complex Banach spaces developed by Sharma and Rebelo. The distinguishing feature of the new calculus is that in this calculus the more general concept of additivity replaces that of linearity in the Frechet calculus. It is proved that the space of continuous additive maps between two complex Banach spaces is the direct sum of the spaces of linear and semilinear maps between the two spaces. The Hahn–Banach theorem and the open mapping theorem which in their standard versions are valid for continuous linear functionals and functions are shown to hold also for the additive case. The concepts of the adjoint of an additive map, of a new kind of orthogonal complement of a subset of a Banach space, and of a balanced additive map in which the norms of the linear and semilinear components are equal are developed. It is then proved that the orthogonal complement of the range of an additive map equals the null space of its adjoint and if the additive map is a functional on a complex Hilbert space and is balanced, then the orthogonal complement of the null space of the functional equals the range of the adjoint. A generalization of the inverse function theorem is proved by using our version of the open mapping theorem and then used to establish the Lagrange multiplier theorem in the new calculus. A number of related results are also proved. The applications of the new calculus to physics are briefly described.

1. INTRODUCTION

The purpose of this paper is to develop further the new calculus on complex Banach spaces introduced by Sharma and Rebelo (1975). This calculus has been used to obtain a variety of results useful in quantum theory by Sharma and Rebelo (1975), Fonte (1979), and Pian and Sharma (1980, 1981).

0020-7748/83/0200-0107$03.00/0

For optimizing functionals on a real Hilbert space, one can use the elegant Frechet calculus, but as Cartan (1971) points out, this calculus cannot be used on the space of nonholomorphic functions on complex normed spaces and all functionals arising from the inner product on a complex Hilbert space are nonholomorphic. On the other hand, the mathematically meaningless rule that a function and its complex conjugate can be varied independently of each other is known to give the correct result consistently and this rule is often used in optimizing functionals on a complex Hilbert space (see, for example, Hirschfelder et al., 1964). Whenever a rule, however arbitrary or inconsistent, consistently gives the correct answer, it points to an underlying mathematical structure which needs to be studied. We believe that the new calculus of Sharma and Rebelo (1975) has isolated the mathematical structure which shows how and why the rule about varying a function and its complex conjugate independently of each other works. Though the new calculus is more complicated than the Frechet calculus on normed spaces, it simplifies considerably the task of optimizing functionals on a complex Hilbert space and in all the cases which arise in quantum theory the new calculus makes it unnecessary either to use the theory of nonholomorphic functions which is considerably more complicated or to use meaningless arbitrary rules.

The main achievements of the present work are the following:

1. We achieve some obvious generalizations of results true for continuous linear maps to continuous additive maps.
2. We are able to show that the space of continuous additive maps between two complex normed spaces is the direct sum of spaces of linear and semilinear maps between the two spaces.
3. We have been able to extend the concept of adjoint to the class of continuous additive maps.
4. We give a new definition of orthogonal complement of a subset of a complex Banach space and prove that the orthogonal complement of the kernel of a balanced (another new concept introduced in the work) additive map equals the range of the adjoint.
5. We use the open mapping theorem to prove a generalized inverse function theorem.
6. We generalize the Lagrange multiplier theorem to the extrema of the moduli of complex functionals.

As far as we know, our proof of the generalized inverse function theorem is a new application of the open mapping theorem and an obvious modification gives a new proof of the result for the linear case.

2. FORMALITIES

The following definitions and notations will be used throughout this work:

Notation 2.1. The rational, the real, and the complex fields will be denoted by $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, respectively.

Notation 2.2. The letters $\mathcal{X}$ and $\mathcal{Y}$ will denote Banach spaces over $\mathbb{C}$.

Definition 2.1. A map f: $\mathcal{X} \to \mathcal{Y}$ is said to be additive if and only if

$$f(x_1+x_2)=f(x_1)+f(x_2), \qquad \forall x_1, x_2 \in \mathcal{X}$$

It should be noted that f is additive implies that $f(qx)=qf(x)$, $\forall q \in \mathbb{Q}$, and f is additive and continuous implies that $f(rx)=rf(x)$, $\forall r \in \mathbb{R}$.

Definition 2.2. A map f: $\mathcal{X} \to \mathcal{Y}$ is said to be *linear* (resp. *semilinear*) if and only if (i) f is additive and (ii) $f(\alpha x)=\alpha f(x)$ [resp. $=\bar{\alpha} f(x)$].

Notation 2.3. $\mathcal{L}(\mathcal{X},\mathcal{Y})$ and $\mathcal{SL}(\mathcal{X},\mathcal{Y})$ will denote the Banach spaces of bounded linear and semilinear functions, respectively, from $\mathcal{X}$ to $\mathcal{Y}$. When $\mathcal{Y}=\mathbb{C}$, the abbreviation $\mathcal{L}(\mathcal{X})$ and $\mathcal{SL}(\mathcal{X})$ will be used for $\mathcal{L}(\mathcal{X},\mathbb{C})$ and $\mathcal{SL}(\mathcal{X},\mathbb{C})$, respectively. It should be noted that if $\mathcal{V}$ is a real normed linear space and f is any additive continuous function from $\mathcal{V}$ to another real normed linear space $\mathcal{W}$, then f is bounded and linear.

Definition 2.3. A map T: $\mathcal{X} \to \mathcal{Y}$ is said to belong to $\mathcal{L}(\mathcal{X},\mathcal{Y}) \oplus \mathcal{SL}(\mathcal{X},\mathcal{Y})$ if and only if T can be written as

$$T={}^{L}T+{}^{S}T$$

with ${}^{L}T \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ and ${}^{S}T \in \mathcal{SL}(\mathcal{X},\mathcal{Y})$. It should be observed that $T \in \mathcal{L}(\mathcal{X},\mathcal{Y}) \oplus \mathcal{SL}(\mathcal{X},\mathcal{Y})$ implies that T is additive and bounded.

Notation 2.4. The notation $\mathcal{A}(\mathcal{X},\mathcal{Y})$ will be used to denote the Banach space of continuous additive functions from $\mathcal{X}$ to $\mathcal{Y}$. It is easily verified that (i) $T \in \mathcal{A}(\mathcal{X},\mathcal{Y}) \Leftrightarrow T$ is additive and bounded and (ii) $T \in \mathcal{A}(\mathcal{X},\mathcal{Y}) \Rightarrow T(0)=0$. (It should be noted that for the second property continuity is not necessary and additivity is enough.)

3. SOME BASIC RESULTS

One of the interesting new findings of this work tells us that $\mathcal{A}(\mathcal{X},\mathcal{Y})$ is in fact the direct sum of $\mathcal{L}(\mathcal{X},\mathcal{Y})$ and $\mathcal{SL}(\mathcal{X},\mathcal{Y})$ and this is the subject matter of our first proposition.

Proposition 3.1. $\mathcal{A}(\mathcal{X},\mathcal{Y})=\mathcal{L}(\mathcal{X},\mathcal{Y})\oplus\mathcal{SL}(\mathcal{X},\mathcal{Y})$.

Proof. In view of the remark under Definition 2.3 it is enough to show that

$$\mathcal{A}(\mathcal{X},\mathcal{Y})\subset\mathcal{L}(\mathcal{X},\mathcal{Y})\oplus\mathcal{SL}(\mathcal{X},\mathcal{Y})$$

Let $T\in\mathcal{A}(\mathcal{X},\mathcal{Y})$. Define F: $\mathcal{X}\to\mathcal{Y}$ by

$$F(x)=(1/2)[T(x)-iT(ix)]$$

Clearly F is additive and continuous since T is and

$$F(ix)=(1/2)[T(ix)+iT(x)]=iF(x)$$

Hence $F\in\mathcal{L}(\mathcal{X},\mathcal{Y})$. Similarly define G: $\mathcal{X}\to\mathcal{Y}$ by

$$G(x)=(1/2)[T(x)+iT(ix)]$$

It is easily verified that $G\in\mathcal{SL}(\mathcal{X},\mathcal{Y})$ and that

$$T=F+G$$

The proof of our proposition, in view of the continuity of T, is now complete. ■

We would remark that the Hamiltonians of quantum mechanics are linear but not continuous. In the preceding proposition the requirement of continuity can be given up if the vector spaces are over the field of the complex rationals or if T in addition to being additive is also linear whenever $\mathcal{X}$ is regarded as a real Banach space. (The underlying set of a complex vector space has the structure of a real vector space with the scalar multiplication which is merely the restriction of the original scalar multiplication to $\mathbb{R}\times\mathcal{X}$; this concept has been used again in the proof of the next proposition).

The preceding proposition shows that a continuous additive function is linear or semilinear or a sum of linear and semilinear functions. It is interesting to note that both $\mathcal{L}(\mathcal{X},\mathcal{Y})$ and $\mathcal{SL}(\mathcal{X},\mathcal{Y})$ are prelinear in the sense of Sharma and Rebelo (1975), but, in general, as is easy to verify, an element of $\mathcal{A}(\mathcal{X},\mathcal{Y})$ need not be prelinear even though it is a sum of two prelinear functions; thus the assertion of Sharma and Rebelo (1975) that the prelinear functions constitute a vector space is in error. Our next proposition shows that the Hahn–Banach theorem is valid also for semilinear functionals.

Proposition 3.2. Let $\mathfrak{M}$ be a subspace of the normed space $\mathfrak{X}$. Let f be a bounded semilinear functional on $\mathfrak{M}$. Then there is a semilinear functional F which is an extension of f to all of $\mathfrak{X}$ and is such that

$$\|F\| = \|f\|$$

Proof. The vectors of $\mathfrak{X}$ form a vector space over $\mathbb{R}$ also, we denote this vector space by $\mathfrak{Z}$. It is evident that as sets

$$\mathfrak{X} = \mathfrak{Z}$$

We denote by $\mathfrak{N}$ the subspace of $\mathfrak{Z}$ spanned by elements of $\mathfrak{M}$. Clearly Ref is a real linear functional on $\mathfrak{N}$. For $x \in \mathfrak{N}$, let

$$f(x) = \mathrm{Re}\, f(x) + \mathrm{Im}\, f(x) = ae^{ir}$$

then since f is semilinear

$$f(e^{ir}x) = a = \mathrm{Re}\, f(e^{ir}x)$$

Hence

$$\|\mathrm{Re}\, f\| = \|f\|$$

Hence by the classical Hahn–Banach theorem Ref has an extension G to all $\mathfrak{Z}$ such that

$$\|G\| = \|\mathrm{Re} f\| = \|f\|$$

As a semilinear functional on $\mathfrak{M}$, f satisfies

$$\mathrm{Re}\, f(im) = \mathrm{Im}\, f(m), \qquad \forall m \in \mathfrak{M}$$

We use this formula to define the semilinear extension F of f on $\mathfrak{X}$ by

$$F(x) = G(x) + iG(ix)$$

Our earlier argument enables us to conclude that

$$\|F\| = \|G\|$$

and there is nothing more to prove. ■

Here we have an example of a deep theorem on linear functionals which has an easy extension to semilinear functionals. Now if we define a norm on $\mathcal{A}(\mathcal{X})$ by

$$\forall T \in \mathcal{A}(\mathcal{X}), \qquad \|T\| = \|{}^L T\| + \|{}^S T\|$$

where $\|{}^L T\|$ is the norm (sup) of the linear functional ${}^L T$ inherited from $\mathcal{L}(\mathcal{X})$ and $\|{}^S T\|$ is the norm (sup) of the semilinear functional ${}^S T$ inherited from $\mathcal{SL}(\mathcal{X})$, then we can extend the Hahn–Banach theorem to $\mathcal{A}(\mathcal{X})$: Proposition 3.2 remains valid if we replace "semilinear" by "additive" throughout in its statement.

We now have some further standard results which have even easier extension to $\mathcal{A}(\mathcal{X})$. Since we need them for proving the Lagrange multiplier theorem, we state these results and indicate briefly how they are proved.

Proposition 3.3. (The open mapping theorem). Let T be a continuous additive map from a Banach space $\mathcal{X}$ onto another Banach space $\mathcal{Y}$. The $T(U)$ is open whenever U is.

Proof. In the proof of the classical open mapping theorem (Dunford and Schwartz, 1957) only additivity has been used although the theorem is stated only for linear functions. Therefore, the same proof applies. ■

> *Corollary 3.3.1.* Let $T \in \mathcal{A}(\mathcal{X}, \mathcal{Y})$ be such that its range $\mathcal{R}(T)$ is a closed subspace of $\mathcal{Y}$. Then there is a real number K such that for each $y \in \mathcal{R}(T)$, there is an $x \in \mathcal{X}$ satisfying
>
> $$Tx = y$$
>
> and
>
> $$\|x\| \leqslant K\|y\|$$

Proof. By the open mapping theorem, the unit ball $\mathcal{S}$ in $\mathcal{X}$ is mapped onto a set $T(\mathcal{S})$, which therefore contains a ball $\mathcal{B}_0(\delta)$ (centered at the origin with radius δ). Let $y \in \mathcal{R}(T)$, then $(\delta y/2\|y\|) \in \mathcal{B}_0(\delta)$ and therefore is the image of x' with $\|x'\| \leqslant 1$. Hence if

$$x = 2\|y\|x'/\delta$$

then

$$Tx = y$$

and

$$\|x\| \leqslant (2/\delta)\|y\|$$

Take

$$K = 2/\delta \qquad \blacksquare$$

Corollary 3.3.2. Let $f \in \mathcal{A}(\mathcal{X}, \mathcal{Y})$ be invertible, then $f^{-1} \in \mathcal{A}(\mathcal{Y}, \mathcal{X})$.

Proof. It is easy to verify that f^{-1} is additive. Since by the open mapping theorem $f = (f^{-1})^{-1}$ maps open sets to open sets f^{-1} is continuous. ■

We next prove a property which is peculiar to elements of $\mathcal{A}(\mathcal{X})$. Every nonzero linear or semilinear functional is surjective. This is not so in the case of elements of $\mathcal{A}(\mathcal{X})$; we now give the necessary and sufficient condition that must be satisfied by an element of $\mathcal{A}(\mathcal{X})$ so that it is surjective.

Proposition 3.4. A necessary and sufficient condition that $f \in \mathcal{A}(\mathcal{X})$ is surjective is that for some $h \in \mathcal{X}$, $|{}^L f(h)| \neq |{}^S f(h)|$, where ${}^L f$ and ${}^S f$ denote the components of f in $\mathcal{L}(\mathcal{X})$ and $\mathcal{SL}(\mathcal{X})$, respectively.

Proof. We may assume that $f \neq 0$. Let $h \in \mathcal{X}$ be such that $|{}^L f(h)| \neq |{}^S f(h)|$. It is easily verified that the equation in z, $z \in \mathbf{C}$, such that

$$f(h_0) \neq 0$$

has a solution for every complex number c. The sufficiency is then proved, since

$$f(zh) = c$$

Conversely, suppose

$$|{}^L f(h)| = |{}^S f(h)|, \qquad \forall h \in \mathcal{X}$$

Choose $h_0 \in \mathcal{X}$ such that

$$f(h_0) \neq 0$$

which implies, in view of our hypothesis, that

$${}^L f(h_0) \neq 0$$

Suppose now that contrary to our proposition f is surjective. Then for every $c \in \mathbf{C}$, there exists an $x \in \mathcal{X}$, such that

$$f(x) = {}^L f(x) + {}^S f(x) = c$$

Since ${}^{L}f \in \mathfrak{L}(\mathcal{X})$ and ${}^{L}f$ is not zero, ${}^{L}f$ is surjective and there exists a $z \in \mathbb{C}$ such that

$$
\begin{aligned}
& {}^{L}f(zh_0) = {}^{L}f(x) \\
\Rightarrow \quad & {}^{L}f(x - zh_0) = 0 \\
\Rightarrow \quad & {}^{S}f(x - zh_0) = 0 \\
\Rightarrow \quad & {}^{S}f(x) = {}^{S}f(zh_0)
\end{aligned}
$$

Hence

$$z\,{}^{L}f(h_0) + \bar{z}\,{}^{S}f(h_0) = c$$

But this equation has a solution for every c if and only if

$$|{}^{L}f(h_0)| \neq |{}^{S}f(h_0)|$$

The contradiction proves the proposition. ■

4. THE ADJOINT

The adjoint plays a crucial role in the theory of Lagrange multipliers. We had to construct a definition of the adjoint for functions belonging to $\mathcal{Q}(\mathcal{X}, \mathcal{Y})$. It will be seen that adjoints of such functions have a much more complicated structure.

Definition 4.1. Let

$$T = {}^{L}T + {}^{S}T \in \mathfrak{L}(\mathcal{X}, \mathcal{Y}) \oplus \mathcal{S}\mathfrak{L}(\mathcal{X}, \mathcal{Y}) = \mathcal{Q}(\mathcal{X}, \mathcal{Y})$$

The adjoint T^* of T is defined to be the map from $\mathcal{Q}(\mathcal{Y})$ to $\mathcal{Q}(\mathcal{X})$ which takes

$$g = {}^{L}g + {}^{S}g \quad \text{to} \quad {}^{L}g\,{}^{L}T + {}^{S}g\,{}^{S}T + {}^{S}g\,{}^{L}T + {}^{L}g\,{}^{S}T$$

(We observe that $gT = T^*g$.)

Definition 4.2. Let $\mathcal{M}$ be a subset of a Banach space $\mathcal{X}$. The orthogonal complement of $\mathcal{M}$ denoted by ${}^{\perp}\mathcal{M}$ is defined to consist of all $f \in \mathcal{Q}(\mathcal{X})$ such that

$$f(m) = 0, \qquad \forall m \in \mathcal{X}$$

[Note that the orthogonal complement of $\mathcal{M}$ is a closed subspace of $\mathcal{Q}(\mathcal{X})$ and $\mathcal{Q}(\mathcal{X})$ contains the dual of $\mathcal{X}$.]

Notation 4.1. In what follows, the range of a function T will be denoted by $\mathcal{R}(T)$ and its kernel by $\mathcal{N}(T)$.

We have now some propositions on the properties of adjoints.

Proposition 4.1. Let $T \in \mathcal{Q}(\mathcal{X}, \mathcal{Y})$. Then ${}^{\perp}[\mathcal{R}(T)] = \mathcal{N}(T^*)$.

Proof. The proof is elementary and similar to the linear case. ■

Proposition 4.2. Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over the complex field. Let $T \in \mathcal{Q}(\mathcal{X}, \mathcal{Y})$ be such that $\mathcal{R}(T)$ is a closed subspace of $\mathcal{Y}$. Then

$$\mathcal{R}(T^*) = {}^{\perp}[\mathcal{N}(T)]$$

Proof. Let $x^* = {}^{L}x^* + {}^{S}x^* \in \mathcal{R}(T^*)$. Then

$$x^* = T^* y^*$$

for some $y^* \in \mathcal{Q}(\mathcal{Y})$. Now

$$x \in \mathcal{N}(T) \Rightarrow x^*(x) = T^* y^*(x) = y^* T(x) = 0$$
$$\Rightarrow x^* \in {}^{\perp}[\mathcal{N}(T)]$$

Hence

$$\mathcal{R}(T^*) \subset {}^{\perp}[\mathcal{N}(T)]$$

Conversely, let $x^* \in {}^{\perp}[\mathcal{N}(T)]$. Define $f\colon \mathcal{R}(T) \to \mathbf{C}$ by

$$f(y) = x^*(x)$$

where

$$y = Tx$$

and x is chosen by Corollary 3.3.1 to satisfy

$$\|x\| \leqslant K \|y\|$$

If there is more than one such x, then since $x^* \in {}^{\perp}[\mathcal{N}(T)]$, $x^*(x)$ has the same value for each such x and, therefore, f is well defined and it is evident that it is additive. It now follows that

$$\|f(y)\| = \|x^*(x)\| \leq K\|x^*\|\|y\|$$

Thus f is bounded in $\mathcal{R}(T)$ and hence by Proposition 3.2 can be extended to all $\mathcal{Y}$. The extension F belongs to $\mathcal{A}(\mathcal{Y})$ and

$$x^* = T^*F$$

Hence ${}^{\perp}[\mathcal{N}(T)] \subset \mathcal{R}(T^*)$ and our proof is complete. ■

When T is a linear operator between two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, T^* according to our definition is a map between $\tilde{\mathcal{H}}_2$ and $\tilde{\mathcal{H}}_1$, the duals of $\mathcal{H}_2$ and $\mathcal{H}_1$, respectively. Since in the case of a Hilbert space the dual is naturally isomorphic to the original space, one can regard T^* as a map from $\mathcal{H}_2$ to $\mathcal{H}_1$ and then our orthogonal complement ${}^{\perp}\mathcal{S}$ of a subset $\mathcal{S}$ of $\mathcal{H}_1$ (or $\mathcal{H}_2$) becomes identical with the usual Hilbert space orthogonal complement $\mathcal{S}^{\perp}$. Further in that case it is easy to see that

$$T^{**} = T$$

With all these identifications $\mathcal{R}(T^*)$ is a subspace of $\mathcal{H}_1$ and $\mathcal{N}(T^*)$ a subspace of $\mathcal{H}_2$. By taking the orthogonal complement of the equality in Proposition 4.1 one gets the equality

$$[\overline{\mathcal{R}(T)}]^{\perp} = [\mathcal{N}(T^*)]$$

where the bar above $\mathcal{R}(T)$ denotes its closure. Replacing T by T^* and by using

$$T^{**} = T$$

one gets two further equalities:

$$[\mathcal{R}(T^*)]^{\perp} = \mathcal{N}(T)$$

and

$$\overline{[\mathcal{R}(T^*)]} = [\mathcal{N}(T)]^{\perp}$$

All these properties are very useful in the study of linear transformations. However, when T is merely additive and continuous, none of the three relations thus obtained are valid and the situation is much more complicated. One source of trouble is that if $\mathcal{S}$ is a subset of $\mathcal{A}(\mathcal{X})$, ${}^{\perp}\mathcal{S}$ defined by

$$^{\perp}\mathcal{S}=\{x: x\in\mathfrak{X}, x^*(x)=0, \forall x^*\in\mathcal{S}\}$$

is not, in general, a subspace of $\mathfrak{X}$. We have investigated the relation between $\mathfrak{R}(T^*)$ and $\mathfrak{N}(T)^{\perp}$ in a particular example of the simple case where $\mathcal{H}_1=\mathcal{H}$ and $\mathcal{H}_2=\mathbf{C}$, so that $T\in\mathfrak{Q}(\mathcal{H})$. This is the subject matter of our longest proposition, but before we can assert it we need a new definition.

Definition 4.3. Let $T\in\mathfrak{Q}(\mathfrak{X},\mathfrak{Y})$. T is said to be *balanced* if

$$\|^{L}T\|=\|^{S}T\|$$

We can now state and prove our next proposition.

Proposition 4.3. Let $\mathcal{H}$ be a complex Hilbert space. Let T be a balanced element of $\mathfrak{Q}(\mathcal{H})$. Let T^* be the adjoint of T. Then

$$^{\perp}[\mathfrak{N}(T)]=\mathfrak{R}(T^*)$$

Proof. Since in the proof of Proposition 4.2 the inclusion

$$\mathfrak{R}(T^*)\subset{}^{\perp}[\mathfrak{N}(T)]$$

was obtained without using the property that $\mathfrak{R}(T)$ is closed, it is valid even in this case and it is enough to prove the inverse inclusion:

$$^{\perp}[\mathfrak{N}(T)]\subset\mathfrak{R}(T^*).$$

Remembering that T is balanced, we can use Riesz representation theorem to write T in the form

$$T=r(\langle\ ,u\rangle+\langle v,\ \rangle)$$

where $r\in\mathbb{R}$, $u,v\in\mathcal{H}$, and

$$\|u\|=\|v\|=1$$

Suppose first that u and v are linearly independent and

$$\mathfrak{U}=\operatorname{Span}\{u,v\}$$

Let P be the orthogonal projection on $\mathfrak{U}$. We first prove that

$$h\in\mathfrak{N}(T)\Rightarrow Ph=\alpha u-\bar{\alpha}v \qquad \text{for some } \alpha\in\mathbf{C}$$

To do this suppose $h \in \mathcal{N}(T)$ and $Ph = \alpha u + \beta v$; then

$$\langle \alpha u + \beta v, u \rangle + \langle v, \alpha u + \beta v \rangle = 0$$
$$\Rightarrow \quad (\alpha + \bar{\beta}) + (\overline{\alpha + \bar{\beta}}) \langle v, u \rangle = 0 \tag{1}$$

Let

$$\langle v, u \rangle = a + ib$$

and

$$\alpha + \bar{\beta} = c + id$$

Then equation (1) can be written

$$(c + id) + (c - id)(a + ib) = 0$$
$$\Rightarrow \quad \left| \begin{array}{ll} c(1+a) + db & = 0 \\ cb & + d(1-a) = 0 \end{array} \right.$$
$$\Rightarrow \quad c = d = 0$$

unless

$$a^2 + b^2 = 1$$

but

$$a^2 + b^2 = |\langle v, u \rangle|^2 < \langle u, u \rangle \langle v, v \rangle = 1$$

We have thus proved that

$$\beta = -\bar{\alpha}$$

Now let

$$x^* = {}^L x + {}^S x \in {}^\perp[\mathcal{N}(T)]$$

that is,

$${}^L x(h) + {}^S x(h) = 0$$

for all h such that

$$Th = 0$$

We can write x^* as

$$x^* = \langle\ , x^1\rangle + \langle x^2, \ \rangle$$

and we shall prove that

$$Px^1 = x^1$$

and

$$Px^2 = x^2$$

Let

$$y^1 = (1-P)x^1$$

and

$$y^2 = (1-P)x^2$$

Clearly y^1 and y^2 belong to $\mathfrak{U}^\perp$ and therefore to $\mathfrak{N}(T)$. Now

$$x^*(y^1) = 0$$

$$\Rightarrow \quad \langle y^1, y^1\rangle + \langle y^2, y^1\rangle = 0$$

and

$$x^*(iy^2) = 0$$

$$\Rightarrow \quad \langle y^2, y^1\rangle - \langle y^2, y^2\rangle = 0$$

Hence

$$\langle y^1, y^1\rangle + \langle y^2, y^2\rangle = 0$$

$$\Rightarrow \quad y^1 = y^2 = 0$$

We can now write

$$x^1 = \alpha_1 u + \beta_1 v$$

and

$$x^2 = \alpha_2 u + \beta_2 v$$

are linearly independent, then

$$x^* \in {}^{\perp}[\mathfrak{N}(T)] \Rightarrow x^* \in \mathfrak{R}(T^*)$$

that is,

$${}^{\perp}[\mathfrak{N}(T)] \subset \mathfrak{R}(T^*)$$

Finally suppose that u and v are linearly dependent, that is,

$$v = \beta u$$
$$\|\beta\| = 1$$

Let

$$\mathcal{V} = \text{Span}\{u\}$$

and let $\mathcal{Q}$ be the orthogonal projection on $\mathcal{V}$. A calculation similar to that in the preceding case shows that if $h \in \mathfrak{N}(T)$, then $\mathcal{Q}h$ can be written as

$$\mathcal{Q}h = \alpha u$$

where α satisfies

$$\alpha + \beta\bar{\alpha} = 0$$

and that if $x^* \in {}^{\perp}[\mathfrak{N}(T)]$, then x^* has the representation

$$x^* = \langle\ , x^1\rangle + \langle x^2, \ \rangle$$

where

$$x^1 = \gamma_1 u$$

and

$$x^2 = \gamma_2 u$$

with

$$\gamma_2/(\bar{\gamma}_1) = \beta$$

We can, therefore, write

$$x^* = T^* g$$

where $g \in \mathcal{L}(\mathbb{C},\mathbb{C})$ is given by

$$g(c)=\bar{\gamma}_1 c$$

Since $\mathcal{L}(\mathbb{C},\mathbb{C}) \subset \mathcal{Q}(\mathbb{C},\mathbb{C})$, the proof of our proposition is complete. ■

5. THE LAGRANGE MULTIPLIER THEOREM

Our main aim in this section is to prove a generalization of the Lagrange multiplier theorem by using the calculus of Sharma and Rebelo (1975). We first recall the basic definition of Sharma and Rebelo (1975) on semidifferentiability and then we prove a series of results culminating in our generalization of the Lagrange multiplier theorem.

Definition 5.1. A function f from a Banach space $\mathcal{X}$ to a Banach space $\mathcal{Y}$ is said to be *semidifferentiable* at a point $x \in \mathcal{X}$, if there exists a function $f_x^{(s)} \in \mathcal{L}(\mathcal{X},\mathcal{Y}) \oplus \mathcal{SL}(\mathcal{X},\mathcal{Y})$ such that

$$\lim_{\|u\| \to 0} \left\| f(x+u)-f(x)-f_x^{(s)}(u) \right\| / \|u\| = 0$$

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 in $\mathcal{X}$, it is said to be *semidifferentiable* in $\mathcal{X}$ and the rule which assigns to each point $x \in \mathcal{X}$ the semiderivative of f at that point is called the semiderivative of f in $\mathcal{X}$ and is denoted by $f^{(s)}$.

Proposition 5.1. (The mean value theorem). Let f: $\mathcal{X} \to \mathcal{Y}$ be semidifferentiable in a neighborhood $\mathcal{U}$ of $x_0 \in \mathcal{X}$. Let $h \in \mathcal{X}$ be such that $x_0 + sh (0 \leq s \leq 1) \in \mathcal{U}$. Then

$$\|f(x_0+h)-f(x_0)\| \leq \sup_{0<\theta<1} \left\| f_{x_0+\theta h}{}^{(s)} \right\| \|h\|$$

Proof. We can regard $\mathcal{Y}$ as a real space. Assume first that

$$y = f(x_0+h) - f(x_0)$$

is a nonzero vector in $\mathcal{Y}$. We define a real function g on the real subspace $\mathrm{Span}\{y\}$ by

$$g(ry) = r\|y\|, \qquad r \in \mathbb{R}$$

Clearly g is linear and

$$\|g\|=1$$

Hence by the Hahn–Banach theorem there exists a real linear continuous function G on $\mathcal{Y}$ such that G extends g and

$$\|G\|=1$$

Now define the real function ρ in $\mathcal{U}$ by

$$\rho(s)=G(f(x_0+sh)), \qquad 0\leqslant s\leqslant 1$$

Then

$$\begin{aligned}\rho'(s)&=\lim_{k\to 0}[\rho(s+k)-\rho(s)]/k\\&=\lim_{k\to 0}G([f(x_0+sh+kh)-f(x_0+sh)]/k)\\&=G\lim_{k\to 0}\{[f(x_0+sh+kh)-f(x_0+sh)]/k\}\\&=G\left(f_{x_0+sh}{}^{(s)}(h)\right)\end{aligned}$$

By the mean value theorem for real functions we have

$$\rho(1)-\rho(0)=\rho'(\theta), \qquad 0<\theta<1$$

that is,

$$G(f(x_0+h)-f(x_0))=G\left(f_{x_0+\theta h}{}^{(s)}(h)\right)$$

Hence

$$\begin{aligned}|G(f(x_0+h)-f(x_0))|&\leqslant\|G\|\left\|f_{x_0+\theta h}{}^{(s)}\right\|\|h\|\\\Rightarrow\|f(x_0+h)-f(x_0)\|&\leqslant\left\|f_{x_0+\theta h}{}^{(s)}\right\|\|h\|\\&\leqslant\sup_{0<\theta<1}\left\|f_{x_0+\theta h}{}^{(s)}\right\|\|h\|\end{aligned}$$

In the case where

$$f(x_0+h)-f(x_0)=0$$

the proposition is obviously true. Our proof is now complete. ■

Definition 5.2. Let T be a continuously semidifferentiable function from an open set $\mathfrak{D}$ in a Banach space $\mathfrak{X}$ into a Banach space $\mathfrak{Y}$, that is, $T^{(s)}$ is defined on the whole of $\mathfrak{D}$ and is continuous. Let $x_0 \in \mathfrak{D}$ be such that $T_{x_0}^{(s)}$ is surjective, then the point x_0 is said to be a *regular point* of the function T.

We now come to the most important new result of this work. This not only generalizes further the generalized inverse function theorem to the case where only semidifferentiability may be available, but also provides a new proof of the standard theorem, that is, for the differentiable case.

Proposition 5.2. (The generalized inverse function theorem). Let x_0 be a regular point of a semidifferentiable map T from a Banach space $\mathfrak{X}$ into a Banach space $\mathfrak{Y}$. Let $T(x_0)=y_0$. Then there is a neighborhood $\mathfrak{U}_{y_0}$ of the point y_0 and a real number K such that the equation

$$T(x)=y$$

has a solution for every $y \in \mathfrak{U}_{y_0}$ and the solution satisfies

$$\|x-x_0\| \leqslant K\|y-y_0\|$$

Proof. We use the open mapping theorem to construct a sequence (h_n), such that $T(x_0+h_n)$ converges to y for y close to y_0. By hypothesis $T_{x_0}^{(s)}$ is surjective. Hence there exists an h_1 such that

$$T_{x_0}^{(s)}(h_1)=y-T(x_0)$$

and by the open mapping theorem h_1 can be chosen in such a way that

$$\|h_1\| \leqslant K\|y-y_0\|$$

Let z_2 be the solution of

$$T_{x_0}^{(s)}(z_2)=y-T(x_0+h_1)$$

such that

$$\|z_2\| \leqslant K\left\|T_{x_0}^{(s)}(z_2)\right\|$$

Let $z_1=h_1$ and let $h_2=z_2+h_1$, then

$$\|h_2-h_1\| \leqslant K\left\|T_{x_0}^{(s)}(h_2-h_1)\right\|$$

We then inductively define z_n to be the solution of

$$T_{x_0}^{(s)}(z_n) = y - T(x_0 + h_{n-1})$$

such that

$$\|h_n - h_{n-1}\| \leq K\left\|T_{x_0}^{(s)}(h_n - h_{n-1})\right\|$$

where

$$h_n = z_n + h_{n-1}$$

We now subtract the equation for $T_{x_0}^{(s)}(z_{n-1})$ from that for $T_{x_0}^{(s)}(z_n)$ to get

$$T_{x_0}^{(s)}(h_2 - h_1) = -\left[T(x_0 + h_1) - T(x_0) - T_{x_0}^{(s)}(h_1)\right] \qquad \text{for } n = 2$$

and

$$T_{x_0}^{(s)}(h_n - h_{n-1}) = -\left[T(x_0 + h_{n-1}) - T(x_0 + h_{n-2}) - T_{x_0}^{(s)}(h_{n-1} - h_{n-2})\right] \qquad \text{for } n > 2.$$

Let $t \in [0,1]$ and let

$$h_{t_2} = th_1$$

and

$$h_{t_n} = th_{n-1} + (1-t)h_{n-2} \quad \text{for } n > 2$$

By the mean value inequality we obtain

$$\left\|T_{x_0}^{(s)}(h_2 - h_1)\right\| \leq \|h_1\| \sup_{t \in]0,1[} \left\|T_{x_0 + h_{t_2}}^{(s)} - T_{x_0}^{(s)}\right\|$$

and

$$\left\|T_{x_0}^{(s)}(h_n - h_{n-1})\right\| \leq \|h_{n-1} - h_{n-2}\| \sup_{t \in]0,1[} \left\|T_{x_0 + h_{t_n}}^{(s)} - T_{x_0}^{(s)}\right\| \qquad \text{for } n > 2.$$

Since $T_{x_0}^{(s)}$ is continuous at x_0, there exists a real number $\delta > 0$ such that

$$\|x - x_0\| < \delta \Rightarrow \|T_x^{(s)} - T_{x_0}^{(s)}\| < 1/2K$$

Now if we take y close enough to y_0, we can assure that

$$\|h_1\| \leqslant K\|y-y_0\| \leqslant \delta/2$$

Then

$$\|h_2-h_1\| \leqslant K\left\|T_{x_0}{}^{(s)}(h_2-h_1)\right\| \leqslant K\|h_1\| \sup_{t\in]0,1[}\left\|T_{x_0+h_{t_2}}{}^{(s)}-T_{x_0}{}^{(s)}\right\|$$
$$\leqslant (1/2)\|h_1\|$$

and

$$\|h_n-h_{n-1}\| \leqslant K\left\|T_{x_0}{}^{(s)}(h_n-h_{n-1})\right\|$$
$$\leqslant K\|h_{n-1}-h_{n-2}\| \sup_{t\in]0,1[}\left\|T_{x_0+h_{t_n}}{}^{(s)}-T_{x_0}{}^{(s)}\right\|$$
$$\leqslant (1/2)\|h_{n-1}-h_{n-2}\|$$

provided

$$\|h_{t_n}\| \leqslant \delta$$

but

$$\|h_{t_n}\| \leqslant t\|h_{n-1}\|+(1-t)\|h_{n-2}\| < \delta$$

if

$$\|h_n\| < \delta \qquad \text{for all } n$$

Now

$$\|h_2\| \leqslant \|h_1\|+\|h_2-h_1\| \leqslant \|h_1\|+(1/2)\|h_1\|$$
$$\leqslant 2\|h_1\| < \delta$$

and, by induction,

$$\|h_n\| \leqslant \|h_1\|+\|h_2-h_1\|+\cdots+\|h_n-h_{n-1}\|$$
$$\leqslant [1+(1/2)+\cdots+(1/2^n)]\|h_1\| \leqslant 2\|h_1\| < \delta \qquad \text{for all } n$$

and in getting the preceding inequality, we have used

$$\|h_n - h_{n-1}\| \leqslant (1/2)\|h_{n-1} - h_{n-2}\| \leqslant (1/2^n)\|h_1\| \qquad \text{for all } n$$

This shows that the sequence (h_n) is a Cauchy sequence which must converge to some vector h, and

$$z_n = h_n - h_{n-1}$$

converges to 0. Then from the equation

$$T_{x_0}^{(s)}(z_n) = y - T(x_0 + h_{n-1})$$

it follows that, in the limit

$$y = T(x_0 + h)$$

Also

$$\|h\| \leqslant 2\|h_1\| \leqslant 2K\|y - y_0\|$$

and the proof of our proposition is finally complete.

We note here that the kernel of an additive map is not necessarily a subspace and, therefore, the proof of the standard theorem (see, for example, Luenberger, 1969) cannot be adapted to be applicable to our case; on the other hand our proof applies equally to the standard case.

Our next proposition completes the path leading to our generalization of the Lagrange multiplier theorem, but we need a definition to state our proposition.

Definition 5.3. Let $T \in \mathcal{Q}(\mathfrak{X}, \mathfrak{Y})$. Let $\mathcal{G}$ be a subspace of $\mathfrak{X}$; then T is said to be *strongly unbalanced on* $\mathcal{G}$ if $\|{}^{L}T(h)\| = \|{}^{S}T(h)\|$ if and only if $h \in \operatorname{Ker} T \cap \mathcal{G}$.

Proposition 5.3. Let $\mathfrak{X}$ be a complex Banach space. Let $\mathfrak{U}$ be an open set containing x_0. Let f be a functional on $\mathfrak{X}$ continuously semidifferentiable on $\mathfrak{U}$ and let H be a mapping from $\mathfrak{X}$ to another complex Banach space $\mathfrak{Y}$ with the following properties:

(i) $f(x)$ restricted to the set $\{x: H(x) = 0\}$ has an extremum at x_0.

(ii) $f_{x_0}^{(s)}$ is strongly unbalanced on $\operatorname{Ker} H_{x_0}^{(s)}$.

(iii) H is continuously semidifferentiable on $\mathfrak{U}$ and $\operatorname{Ker} H_{x_0}^{(s)}$ is a subspace.

(iv) x_0 is a regular point of both H and f.

Then

$$f_{x_0}^{(s)}(h)=0$$

for all h satisfying

$$H_{x_0}^{(s)}(h)=0$$

Proof. We define a mapping T: $\mathcal{X} \to \mathbf{C} \oplus \mathcal{Y}$ by

$$T(x)=(f(x), H(x))$$

Suppose contrary to our assertion there exists an $h \in \mathcal{X}$ such that

$$H_{x_0}^{(s)}(h)=0$$

but

$$f_{x_0}^{(s)}(h) \neq 0$$

From the postulated properties of f and H it is now easy to verify that

$$T_{x_0}^{(s)}=\left(f_{x_0}^{(s)}, H_{x_0}^{(s)}\right)$$

is a surjective map from $\mathcal{X}$ to $\mathbf{C} \oplus \mathcal{Y}$. Then by Proposition 5.2 there exists a neighborhood $\mathcal{V}$ of $T(x_0)$ such that the equation

$$T(x)=y$$

has a solution for every $y \in \mathcal{V}$. In particular one can choose a positive real number r such that

$$T(x)=(f(x_0) \pm \delta, 0)$$

has a solution for every $\delta < r$. This contradicts that $|f(x_0)|$ is an extreme value of $|f|$. This completes the proof. ■

We would remark that the requirements that $f_{x_0}^{(s)}$ be strongly unbalanced and that $\operatorname{Ker} H_{x_0}^{(s)}$ be a subspace are unnecessary when f is a real functional on a complex Banach space and furthermore the requirement that f' be regular is unnecessary when f is a real functional on a real Banach space. All these requirements are to assure the surjectivity of the semiderivative or the derivative and all real functionals are either zero or surjective.

We would remark further that the main assertion of Proposition 5.3 can be restated as

$$f_{x_0}^{(s)} \in {}^{\perp}\left[\mathfrak{N}\left(H_{x_0}^{(s)}\right)\right]$$

We have now proved everything necessary to be able to assert our main result without any further proof.

Proposition 5.4. (The generalized Lagrange multiplier theorem). Let f and H be as in Proposition 5.3. Then there exists an element $\lambda \in \mathcal{Q}(\mathcal{Y})$ such that the functional

$$f_{x_0}^{(s)} + \lambda \circ H_{x_0}^{(s)} = 0$$

We finally remark that the same method will yield the corresponding results for the extrema of real functionals on real or complex Banach spaces. Physical applications of these results are being published elsewhere (Pian and Sharma, 1981).

6. APPLICATIONS TO PHYSICS

Sharma and Rebelo (1975) have discussed the importance of optimization in science and industry. It is well known that a good proportion of laws of physics can be obtained by a variational principle and a variational principle is often the only means of placing bounds on errors in approximate calculations of physical quantities. The discovery of Frechet calculus led to a substantial simplification of the calculus of variations and it is now possible to provide short, rigorous, and illuminating proofs of classical results which were originally established by methods which were not only a little obscure but also lacked rigor. Theoretical physicists have so far been unable to benefit from this development because the Frechet calculus could not be used for optimization of functionals and functions which are partly semilinear and semilinearity is a necessary attribute of an inner product on a complex Hilbert space which is the home of much of theoretical physics at the present time. Physicists had to rely on classical methods which included such meaningless stratagems as varying a function and its complex conjugate independently of each other. The new calculus makes this unnecessary and brings the modern development into a form which can be used for optimization of functionals and functions defined on a complex Hilbert space. The new calculus has been used (Sharma and Rebelo, 1975) to obtain bounds on a variety of quantum mechanical sums; these sums include many of great practical importance such as the second-order energies of stationary systems, atomic polarizabilities, and so on. Fonte (1979) has used the calculus to provide a rigorous modern proof of the Ritz–Rayleigh bounds on eigenvalues of a Hamiltonian operator in quantum mechanics. The new calculus has also been used to provide simple, rigorous proofs of Kato's lemma on Schwinger's variational principle for the scattering phase (Pian and Sharma, 1980), of a generalized Brillouin theorem (Pian and Sharma, 1981) and of the Hartree–Fock equations for two electron atoms (Pian, 1981). Further work is in progress and will be reported in due course.

ACKNOWLEDGMENT

The first author gratefully acknowledges the award of a University Research Studentship from the University of London during the tenure of which this work was done.

REFERENCES

Cartan, H. (1971). *Differential Calculus*. Hermann, Paris.

Dunford, N., and Schwartz, J. T. (1957). *Linear Operators Part I*. Interscience, New York.

Fonte, G. (1979). *Nuovo Cimento*, **49B**, 200–219.

Hirschfelder, J. O., Brown, W. B., and Epstein, S. T. (1964). *Advances in Quantum Chemistry*, **1**, 255–374.

Luenberger, D. G. (1969). *Optimization by Vector Space Methods*. John Wiley, New York.

Pian, J. (1981). Ph.D. thesis, London University.

Pian, J., and Sharma, C. S. (1980). *Physics Letters*, **76A**, 365–366.

Pian, J., and Sharma, C. S. (1981). *Journal of Physics A (London)*, **14**, 1261–1270.

Sharma, C. S., and Rebelo, I. (1975). *International Journal of Theoretical Physics*, **13**, 323–336.