A GENERAL FORMULATION OF VARIATIONAL PRINCIPLES

Ismael Herrera*

* Consultant, Institute of Engineering, National University of Mexico
Research Professor, Institute of Geophysics, National University of Mexico

1. INTRODUCTION 1

2. PRELIMINARY NOTIONS AND NOTATIONS 7

3. VARIATIONAL PRINCIPLES 15

4. EXTREMAL PRINCIPLES 16

5. DUAL VARIATIONAL PRINCIPLES 22

6. HAMILTON PRINCIPLES 30

7. APPLICATIONS 32

APPENDIX 41

REFERENCES 44

1. INTRODUCTION

This article presents a general formulation of variational principles. By a variational principle is understood and assertion stating that the variation or derivative of some functional vanishes if and only if a given equation is fulfilled. An extremal principle is one which establishes the equivalence between an equation and the fact that some functional attains an extremum value, i.e. either a maximum or a minimum.

When the functional is differentiable, a necessary condition for the existence of a maximum or a minimum is the vanishing of its variation or derivative. This condition is not sufficient and consequently the class made of all extremal principles is a proper subset of the class made of all variational principles.

The theoretical foundations of variational methods rest on the theory of differentiation on Hilbert spaces, or more generally on Banach spaces {1}; more specifically on the results of that theory related to the notion of potential operators {2}. Essentially, one can say that a sufficient condition for an operator equation to admit a variational formulation is that the operator be potential and this is the case if and only if its derivative is a symmetric bilinear functional.

When the operator is not potential it is always possible, at least in principle, to transform the problem into an equivalent one for which the operator involved is potential. Although it has been claimed that this is the key by which we may obtain a very wide class of variational principles, applicable to almost any equation {3}, its limited practical value due to the difficulties involved in finding such transformation of the problem must be kept in mind. However, this approach has been used successfully to treat initial value problems {4, 5}.

In many applications a problem can be formulated variationally in two different ways related to each other, and the solution is characterized by both a maximum and a minimum principle. The maximum and minimum values of the respective functionals are the same. In this case the principles are said to be dual, complementary, or reciprocal {6 - 11}.

The value of dual variational principles is great, because the difference between the values of both functionals for a couple of trial functions can be used as a measure of the accuracy of approximate solutions. In many cases the significance of dual variational principles is further enhanced because the functionals involved have some relevant meaning in a field of application; e.g. the energy in some physical applications.

Some complementary variational principles {10} were already known a long time ago[+], but it was not until 1929 that Friedrichs gave the first systematic account of them. Courent and Hilbert {12}, based their treatment on Friedrich's work. Later Prager and Synge developed the hypercicle method {13}, Díaz and Weinstein developed an approach suitable for

---

+ For example, Castigliano apparently discovered in 1873 {12} that the minimum potential energy theorem for an elastic body has a dual related to "complementary" energy.

linear problems {14} and Kato and Fujita {15, 16} studied the operator decomposition $L = T^*T$. Applications of the principles have been abundant and accounts of them are available {6 - 11}.

The work of Noble and Sewell on the subject is especially relevant. According to them {11}, the class of dual variational principles is made up of generalized Lagrange principles and generalized Hamilton principles. Lagrange principles apply to a system of equations of the form

$$T^*u = y \tag{1.1.a}$$

$$T \frac{\partial Y}{\partial y} = - \frac{\partial Y}{\partial u} \tag{1.1.b}$$

where T is a linear operator, T^* its adjoint and Y a nonlinear functional. Hamilton principles apply to the system of equations:

$$\begin{aligned} T^*u &= \frac{\partial X}{\partial x} \\ Tx &= \frac{\partial X}{\partial u} \end{aligned} \tag{1.2.a}$$

Here X is a nonlinear functional. The extremal principles are granted if some functionals defined in terms of Y and X, have properties connected with the notions of convex, concave and saddle-shaped functionals.

It seems, however, there are some important points that must be clarified further. First, there are many definitions of a derivative or variation of a functional or operator {1} and flexibility in applications of the theory depends on which is used. Vainberg {2} observed that the Frechet derivative is too restrictive and consequently used the Gateaux derivative in his treatment of the subject. However, use of the Gateaux derivative requires the space to be a Banach space or at least a normed space, in which the derivative must be a continuous linear operator. When an equation is set within the framework of variational calculus, the operator involved in the equation is required to be the derivative of some functional and consequently it must be continuous with respect to the given norm. This requirement complicates applications of the theory, especially when dealing with boundary conditions. Thus, in many cases very complicated or cumbersome norms and inner products have to be introduced {9 - 11}. It seems that the introduction of such norms or inner products is irrelevant, at least as regards the formulation of variational principles. In this work, additive Gateaux variations in the sense of Nashed {1}, will be used. The theory so obtained has great flexibility, because it can be applied when the operators considered are defined in a linear space which is not equiped with an inner product or a norm.

There is at least one other point that must be clarified. Noble and Sewel {11} state in the introduction to their account of dual variational principles:

"Finally we remark that this paper is in no sense intended to be encyclopedic or exhaustive. Instead it is intended to convey an approach to dual extremum principles which the reader may find fruitful in developing explicit consequences for himself. The examples demostrate a worthwhile range of applicability, without justifying any claim of universality".

It is quite clear that Noble and Sewell's work is indeed very worthwhile. but it would also be worthwhile to develop a general scheme which allowed us to place dual variational principles in due perspective within the general framework of the calculus of variations. The relation between Lagrange and Hamilton principles is well established {11}, they can be ob tained from earch other by Legendre transformations. On the other hand, Rall {17} has developed a model for operators defined on a Hilbert space, which makes it possible to put dual variational principles of the generalized Hamilton type into a functional-analytical framework. However, the main point that remains to be clarified in order to place dual variational principles in perspective within the general framework of calculus of variations is concerned with the establishment of clear relations between dual variational principles and other kinds of variational principles, and Rall's work is not concerned with this question.

In this article a general formulation of variational principles is developed, including dual variational principles. In this manner the connections between dual variational principles and other kinds of principles can be better understood. Also, a new class of variational principles is obtained.

First, the basic notions and notations that will be used are introduced. As mentioned previously, the theory is developed for operators defined on a linear space in which neither an inner product nor a norm is assumed to be present. The operators are supposed to be functional valued and to possess additive Gateaux variations {1}. To keep the formulation of the theory as simple as possible, and in this manner emphasize its most relevant aspects, all the operators are assumed to be defined in the whole linear space considered. For operators which are only defined in a subset of a linear space, the development of the corresponding modifications of the theory is straight forward.

Potential operators are discussed in this setting and the symmetry condition is established. When the operator is linear, this condition is equivalent to requiring that the operator itself be symmetric. Extremal principles are obtained for positive or alternatively negative operators.

These results for linear operators are generalized to nonlinear ones introducing the notion of convex and concave functionals.

In general a symmetric linear operator L, is neither positive nor negative. However, the linear space D can frequently be decomposed into two subspaces D_+ and D_- such that L is positive in D_+ while negative in D_- and $D = D_+ + D_-$. For example, if L is a self-adjoint (continuous) linear operator of a real Hilbert space into itself, such a decomposition is guaranteed by the spectral theorem. When this decomposition is available, it is shown that the variational equation corresponding to the equation associated with the operator L can be split into two. To this end partial variations associated with D_+ and D_- are introduced. This pair of equations is used to define a new class of dual variational principles. These variational principles are generalized to nonlinear equations introducing the notion of saddle functionals. It is then shown how Hamilton dual variational principles can be derived from them. In the last section some applications are carried out to illustrate the theory. In particular some dual variational principles for initial value problems of the heat equation are obtained; apparently they were not available previously.

To summarize, the formulation of variational principles presented in this paper has the following features:

i) It is developed in terms of operators defined on a linear space (without inner product or norm), which are only required to possess an additive Gateaux variation (not necessarily continuous). This formulation is very suitable to the handling of boundary conditions in a direct and rigorous manner {5}.

ii) A new general kind of dual variational principles is introduced. It constitutes a link which makes a unified formulation of the different classes of variational principles so far known, possible.

iii) This new class of dual variational principles introduced, yields valuable results in specific applications. To illustrate this point, dual variational principles for initial value problems of the heat equation are derived.

iv) The theory is based on very simple mathematical facts. This feature has didactic value, because the theory can therefore can be grasped by people with only a limited background in Functional Analysis.

It is convenient to recall that the terminology applied to different kinds of derivatives of functionals and operators changes very much with different authors. In this paper the terminology presented by Nashed {1} in his account, is followed to a large extent.

2. PRELIMINARY NOTIONS AND NOTATIONS

All the linear spaces to be considered will be defined on the field of real numbers R^1. Given two linear spaces D_1 and D_2, the set $D_1 \times D_2$ is defined by:

$$D_1 \times D_2 = \{(x_1, x_2) \mid x_1 \varepsilon D_1 \text{ and } x_2 \varepsilon D_2\} \quad (2.1)$$

The set $D_1 \times D_2$ becomes a linear space if the operations of sum and multiplication by a real number are defined by:

$$x + y = (x_1 + y_1, x_2 + y_2) \quad (2.2.a)$$

$$ax = (ax_1, ax_2) \quad (2.2.b)$$

whenever $x = (x_1, x_2)$, $y = (y_1, y_2) \varepsilon D_1 \times D_2$ and $a \varepsilon R^1$. In view of the fact that the product between linear spaces is an associative operation, the product $D_1 x \ldots x D_n$ between n linear spaces $D_1, \ldots, D_n$, is well defined.

If $D = D_1, \ldots, D = D_n$, the notation

$$D^n = D_1 \times \ldots \times D_n \tag{2.3}$$

will be used.

Each of the spaces D_j, $j = 1,\ldots,n$, is isomorphic to the linear subspace:

$$\hat{D}_j = \{(0,\ldots,0, x_j, 0,\ldots,0)\} \subset D_1 \times \ldots \times D_n \tag{2.4}$$

Observe in addition that

$$\hat{D}_j \cap \hat{D}_k = \{0 \varepsilon D_1 \times \ldots \times D_n\} \text{ , whenever } j \neq k \tag{2.5.a}$$

$$D_1 \times \ldots \times D_n = \hat{D}_1 + \hat{D}_2 + \ldots + \hat{D}_n \tag{2.5.b}$$

These properties will be used in the sequel. On the other hand, D_1, $D_2 \subset D$ are called a decomposition of D if they are independent and span D.

Given two linear spaces D and D', an operator

$$P : D \to D' \tag{2.6}$$

is said to be linear if it is additive and homogeneous+. The value of a functional:

$$\alpha : D^n \to R^1 \tag{2.7}$$

at an element $x = (x_1,\dots,x_n) \in D^n$ will be represented by $\alpha(x_1,\dots,x_n)$. Such a functional is said to be n-linear in D if $\alpha(x_1,\dots,x_n)$ is linear in each x_j when all the other arguments are kept fixed. For n-linear functionals the notation

$$\langle\alpha, x_1,\dots,x_n\rangle = \alpha(x_1,\dots,x_n) \tag{2.8}$$

will be used. It must be recalled that no inner product will be used in this paper and that the notation introduced by equation 2.8 does not imply that an inner product in the linear space D is assumed to be present.

The set of n-linear functionals will be represented by D^{n*}. When $n = 1$, this is the dual space of D that will be represented by D^*. In addition, the notation

$$D^{\circ *} = R^1 \tag{2.9}$$

will be adopted. With the usual definitions of addition and multiplication of functions, by a real, the set D^{n*} is a linear space for every $n = 0, 1,\dots$

In this article, functional valued operators

$$P : D \to D^{n*} \tag{2.10}$$

+ For some authors such as Vainberg {2} the term linear operator includes the hypothesis of continuity. Observe that this is not required in the sense the term is used in this paper.

play an important role. Notice that when $n = 0$, the operator P given in equation 2.10 is by virtue of equation 2.9 a functional. When $n = 1$, for every $x, y \varepsilon D$, $\langle P(x), y\rangle \varepsilon R^1$. As a matter of fact $\langle P(x), y\rangle$ defines a functional on D^2 linear in y. Obviously, every functional on D^2 which is linear with respect to the second argument defines an operator

$$P : D \to D^*$$

In particular, if $P = L$ is linear, the corresponding functional is bilinear in D. Therefore, given $L : D \to D^*$, the adjoint operator $L^* : D \to D^*$ can be defined by:

$$\langle L^* x, y\rangle = \langle Ly, x\rangle \tag{2.11}$$

which holds for every $x, y\varepsilon D$.

For operators of the type equation 2.10 the notion of continuity can be introduced without a toplogy in D.

2.1 Definition

Let $P : D \to D^{n*}$. Given $x\varepsilon D$, P is said to be bidimensionally continuous at x, if for every $y, z, \xi_1,\ldots,\xi_n \varepsilon D$ the function $f(\eta,\lambda)$ of the two real variables η and λ defined by:

$$f(\eta, \lambda) = \langle P(x + \eta y + \lambda z), \xi_1,\ldots,\xi_n\rangle \tag{2.12}$$

is continuous at $\eta = \lambda = 0$.

Given $x, y, \xi_1, \ldots, \xi_n \varepsilon D$ let g be a function of the real variable t defined by:

$$g(t) = \langle P(x + ty), \xi_1, \ldots, \xi_n \rangle \qquad (2.13)$$

Fixed $x \varepsilon D$, if the derivative $g'(0)$ exists for every $y, \xi_1, \ldots, \xi_n \varepsilon D$, it is necessarily linear in each one of $\xi_1, \ldots, \xi_n$, so that

$$g'(0) = \langle \nabla P(x, y), \xi_1, \ldots, \xi_n \rangle \qquad (2.14)$$

where $\nabla P(x, y) \varepsilon D^{n*}$. In addition, $g'(0)$ is also necessarily homogeneous in y. If $g'(0)$ is additive in y, one can write:

$$g'(0) = \langle P'(x), y, \xi_1, \ldots, \xi_n \rangle \qquad (2.15)$$

where $P'(x) \varepsilon D^{(n + 1)*}$.

2.2 *Definition*

Given $x \varepsilon D$, if a functional $P'(x) \varepsilon D^{(n + 1)*}$ exists such that with $g(t)$ given by equation 2.13, equation 2.15 holds for every $y, \xi_1, \ldots, \xi_n \varepsilon D$, then it is said that the Gateaux variation of P exists and is additive[+] at x.

For the additive Gateaux variation, the alternative notation

$$\frac{dP}{dx}(x) = p'(x) \qquad (2.16)$$

+ Observe that $P'(x)$ is only assumed to be additive but not continuous.

will be used. The nth-order additive Gateaux variation is defined inductively and will be denoted by $P^{(n)}(x)\varepsilon D^{n*}$.

When the linear space D is the direct sum of two independent subspaces D_1 and D_2, then for every $y\varepsilon D$ there exist $y_1\varepsilon D_1$ and $y_2\varepsilon D_2$ uniquely defined such that

$$y = y_1 + y_2 \tag{2.17}$$

In this case, if the additive Gateaux variation exists at $x\varepsilon D$, , two functionals $P_{,1}(x)\varepsilon D^{(n+1)*}$ and $P_{,2}(x)\varepsilon D^{(n+1)*}$ can be defined by:

$$<P_{,1}(x),\ y,\ \xi_1,\dots,\xi_n> = <P'(x),\ y_1,\ \xi_1,\dots,\xi_n> \tag{2.18.a}$$

$$<P_{,2}(x),\ y,\ \xi_1,\dots,\xi_n> = <P'(x),\ y_2,\ \xi_1,\dots,\xi_n> \tag{2.18.b}$$

for every $y,\xi_1,\dots,\xi_n\,\varepsilon D$. The functionals $P_{,1}(x)$ and $P_{,2}(x)$ will be called the partial Gateaux variations of P in the subspace D_1 and in the subspace D_2, respectively. Observe that

$$P'(x) = P_{,1}(x) + P_{,2}(x) \tag{2.19}$$

The alternative notation

$$\frac{\partial P}{\partial x_1}(x) = P_{,1}(x) \tag{2.20.a}$$

$$\frac{\partial P}{\partial x_2}(x) = P_{,2}(x) \tag{2.20.b}$$

will be used. Higher order partial variations are defined inductively.

The following theorems and definition are quoted for further reference.

2.3 *Taylor's theorem*

Let

i) $\psi : D \to R^1$

ii) ψ have additive Gateaux variations up to order $n + 1$ in D

iii) the operator $\psi^{(n+1)} : D \to D^{(n+1)*}$ be bidimensionally continuous in D.

Then given $x, y \varepsilon D$, there exists ξ belonging to the segment that joins x with y such that:

$$\psi(y) = \psi(x) + \langle\psi'(x), y - x\rangle + \ldots + \frac{1}{n!}\langle\psi^{(n)}(x), y - x, \ldots, y - x\rangle$$

$$+ \frac{1}{(n+1)!}\langle\psi^{(n+1)}(\xi), y - x, \ldots, y - x\rangle \qquad (2.21)$$

2.4 *Definition*

An operator $P : D \to D^*$ is said to be potential if there exists a functional

$$\psi : D \to D^{o*} = R^1$$

such that

$$\frac{d\psi}{dx}(x) = P(x) \qquad (2.22)$$

for every $x \varepsilon D$. In such a case the functional ψ is said to be a potential of P.

2.5 Theorem

Let

i) $P : D \to D^*$;

ii) P have an additive Gateaux variation $P'(x)$ for every $x \varepsilon D$;

iii) the bilinear functional $P'(x)$ be bidimensionally continuous for each $x \varepsilon D$.

Then a necessary and sufficient condition for P to be a potential operator is that for every $x \varepsilon D$ the bilinear functional $P'(x)$ be symmetric; i.e.

$$\langle P'(x), y, z \rangle = \langle P'(x), z, y \rangle \tag{2.23}$$

for every $y, z \varepsilon D$.

Proofs.- The proof of Theorem 2.3 follows from a straight forward application of Taylor's theorem with residue for functions of one variable. The proof of Theorem 2.5 can be carried out in a similar manner to that suggested by Vainberg {2, pp. 56}; the details are given in the Appendix.

Of special interest is the case when the operator $P = L$ is linear.

2.6 Corollary

Let $L : D \to D^*$ be a linear operator. Then L is potential if and only if L is self-adjoint.

Proof.- It is an obvious consequence of Theorem 2.5.

3. VARIATIONAL PRINCIPLES

The general problem to be considered consists in finding $x \varepsilon D$ such that

$$P(x) = f \qquad (3.1)$$

where P is a functional valued operator $P : D \to D^*$, $f \varepsilon D^*$ and D is a linear space.

By a variational principle is understood an assertion that establishes that the variation of a functional vanishes at a point if and only if the equation 3.1 is satisfied. A sufficient condition for the construction of variational principles is formulated in the theorem that follows.

3.1 Theorem

Let P be potential and $\psi : D \to R^1$ be a potential of P. Define $\Omega : D \to R^1$ by

$$\Omega(x) = \psi(x) - \langle f, x \rangle \qquad (3.2)$$

Then, $x_o \varepsilon D$ is a solution of equation 3.1 if and only if

$$\frac{d\Omega}{dx}(x_o) = 0 \qquad (3.3)$$

Proof.- The proof follows from the fact that:

$$\frac{d\Omega}{dx}(x) = P(x) - f \qquad (3.4)$$

In the case of potential linear operators

$$\psi(x) = \frac{1}{2} <Lx, x> \tag{3.5}$$

and consequently

$$\Omega(x) = \frac{1}{2} <Lx, x> - <f, x> \tag{3.6}$$

4. EXTREMAL PRINCIPLES

In what follows a point $x_o \varepsilon D$ will be said to be a maximum of a functional Ω if

$$\Omega(x_o) \geq \Omega(x) \tag{4.1}$$

for every $x \varepsilon D$. The maximum will be said to be strict if the inequality holds whenever $x \neq x_o$. Minimum and strict minimum are defined in a similar manner.

By a maximum principle is understood an assertion establishing that a certain functional attains a maximum at a point if and only if equation 3.1 is satisfied there. An analogous definition is given for a minimum principle. An extremal principle is one which is either a maximum or a minimum principle.

As is well-known, under very general conditions any extremal principle is a variational principle.

4.1 Theorem

Let $\Omega : D \to R^1$ possess an additive Gateaux variation at $x_o \varepsilon D$. If Ω attains either a maximum or a minimum at x_o, then

$$\frac{d\Omega}{dx}(x_o) = 0 \tag{4.2}$$

Proof.- Because, for every $y\varepsilon D$:

$$<\frac{d\Omega}{dx}(x_o),\ y> = \lim_{t \to 0} \frac{\Omega(x_o + ty) - \Omega(x_o)}{t} \leq 0 \tag{4.3.a}$$

and

$$-<\frac{d\Omega}{dx}(x_o),\ y> = \lim_{t \to 0} \frac{\Omega(x_o - ty) - \Omega(x_o)}{t} \leq 0 \tag{4.3.b}$$

Thus

$$<\frac{d\Omega}{dx}(x_o),\ y> = 0 \tag{4.4}$$

for every $y\varepsilon D$.

A linear operator $L : D \to D^*$ is said to be positive if

$$<Lx,\ x> \geq 0 \tag{4.5}$$

for every $x\varepsilon D$. It is strictly positive if the inequality holds in equation 4.5 whenever $x \neq 0$. For positive operators it is easy to formulate minimum principles.

4.2 Theorem

Let the operator $P = L : D \rightarrow D^*$ be linear and potential. Assume further that L is positive. Then $\Omega : D \rightarrow R^1$ as given by equation 3.6 attains a minimum at $x_o \varepsilon D$ if and only if x_o satisfies equation 3.1.

Proof.- First, a convenient notation is introduced. For any x_-, $x_+ \varepsilon D$, the following definitions are given

$$\Delta\Omega = \Omega(x_+) - \Omega(x_-) \qquad (4.6.a)$$

$$\Delta x = x_+ - x_- \qquad (4.6.b)$$

$$\frac{d\Omega}{dx_+} = \frac{d\Omega}{dx}(x_+) \qquad (4.6.c)$$

$$\frac{d\Omega}{dx_-} = \frac{d}{dx}(x_-) \qquad (4.6.d)$$

Then, set $x_- = x_o$ and observe that

$$0 \leq \frac{1}{2} <L\Delta x, x> = \Delta\Omega - <\frac{d\Omega}{dx_-}, \Delta x> \qquad (4.7)$$

But

$$\frac{d\Omega}{dx_-} = Lx_o - f \qquad (4.8)$$

Thus, if x_o satisfies equation 3.1 then

$$\Omega(x_+) \geq \Omega(x_o) \tag{4.9}$$

for any $x_+ \varepsilon D$ and consequently x_o is a minimum.

Conversely, if x_o is a minimum equation 3.1 is satisfied by virtue of Theorem 4.1.

There are two obvious corollaries to Theorem 4.2; namely:

4.3 Corollary

If in Theorem 4.2, L is assumed to be strictly positive then Ω attains strict minimum at $x_o \varepsilon D$ if and only if x_o satisfies equation 3.1.

4.4 Corollary

If L is strictly positive, then equation 3.1 possesses at most one solution.

The extension of the results of Theorem 4.2 to non-linear operators is straight forward if use is made of the notion of convex and concave functionals. To introduce these concepts, it is convenient to retain the notation used in the proof of Theorem 4.2.

4.5 Definition

A functional $\Omega : D \to R^1$ possessing an additive Gateaux variation for every $x \varepsilon D$, is said to be convex if

$$\Delta\Omega - \langle \frac{d\Omega}{dx\_}, \Delta x \rangle \geq 0 \tag{4.10}$$

for every x_+, $x_- \varepsilon D$. It is strictly convex if the strict inequality holds whenever $x_+ \neq x_-$. In addition Ω is concave or strictly concave if $-\Omega$ is convex or strictly convex respectively.

It is convenient to observe that the inequality 4.10 can be replaced by

$$\Delta\Omega - \langle \frac{d\Omega}{dx_+} , \Delta x \rangle \leq 0 \tag{4.11}$$

Indeed, equation 4.10 and 4.11 are obtained from each other, interchanging the roles of x_+ and x_-.

4.6 Theorem

Let P and Ω be as in Theorem 3.1. Assume Ω to be convex. Then $x_o \varepsilon D$ satisfies equation 3.1 if and only if x_o is a minimum of Ω.

Proof.- The inequality equation 4.10 is the same as inequality equation 4.7. Thus, the proof can be carried out along the same lines as that of Theorem 4.2.

The results contained in corollaries 4.2 and 4.3 follow from Theorem 4.6, in the case when Ω is strictly convex. It is also convenient to observe that Ω is convex if and only if Ψ is convex, because:

$$\Delta\Omega - \langle \frac{\partial\Omega}{\partial x_-} , \Delta x \rangle = \Delta\Psi - \langle \frac{d\Psi}{dx_-} , \Delta x \rangle \tag{4.12}$$

The concept of a convex functional is a natural generalization of the concept of positive operator and therefore Theorem 4.6 is natural generalization of Theorem 4.2. This can be better appreciated by recalling the theorem that follows.

4.7 Theorem

Let $\Omega : D \to R^1$ be such that:

i) Ω has second order additive Gateaux variation at every $x \epsilon D$;

ii) Let $\Omega'' : D \to D^{2*}$ be bidimensionally continuous in D.

Then a necessary and sufficient condition for Ω to be convex is that for every $x \epsilon D$, $\Omega''(x) \epsilon D^{2*}$ be positive.

Proof.- In view of Taylor's Theorem 2.3 for every x_- , $x_+ \epsilon D$, we have:

$$\Delta\Omega - \langle \frac{d\Omega}{dx_+} , \Delta x \rangle = \frac{1}{2} \langle \Omega''(\xi), \Delta x, \Delta x \rangle \qquad (4.13)$$

where ξ belongs to the segment joining x_- with x_+. Therefore, if $\Omega''(\xi)$ is positive for every $\xi \epsilon D$, Ω is convex.

To prove the converse, assume that for some $y \epsilon D$, $\Omega''(y)$ is not a positive bilinear functional. Then, there exists some $z \epsilon D$ such that $\langle \Omega''(y), z, z \rangle < 0$. The function of the real variable λ given by $\langle \Omega''(y + \lambda z), z, z \rangle$ is continuous and consequently there is some real $\lambda_0 < 0$ with the property that for any real λ_1 such that $|\lambda_1| < \lambda_0$:

$$\langle \Omega''(y + \lambda_1 z), z, z \rangle < 0 \qquad (4.14)$$

Chose a λ_1 satisfying this condition and set $x_- = y - \lambda_1 z$; $x_+ = y + \lambda_1 z$. Then, by Taylor's Theorem 2.3, for some real λ_2 which satisfies the condition $|\lambda_2| \leq |\lambda_1|$, we have

$$\Delta\Omega - \langle\frac{d\Omega}{dx_+}, \Delta x\rangle = 2\lambda_1^2 \langle\Omega''(y + \lambda_2 z), z, z\rangle < 0 \tag{4.15}$$

Thus, Ω is not convex and the Theorem is proved.

Evidently, the concepts maximum-minimum, covex concave, positive-negative constitute dual couples and it can be seen that the theory developed in this section remains valid if each of these concepts is interchanged with each of its duals.

5. DUAL VARIATIONAL PRINCIPLES

Theorem 4.2 by establishing extremal principles for positive linear operators, has restricted applicability because there are many linear operators of interest which are neither positive nor negative. A theorem will now be formulated that has a larger range of applicability.

Consider a linear and potential operator $P = L : D \to D^*$. Assume there exists a decomposition D_1, D_2 of D, such that

$$\Psi(x) = \frac{1}{2}\langle Lx, x\rangle \geq 0 \tag{5.1.a}$$

whenever $x \varepsilon D_1$ and

$$\Psi(x) = \frac{1}{2}\langle Lx, x\rangle \leq 0 \tag{5.1.b}$$

whenever $x \varepsilon D_2$. On the other hand, the definition of partial Gateaux variations equations 2.18,can be applied because D_1, D_2 are a decomposition of D. When this is done, equation 2.19 implies that

$$\frac{d\Omega}{dx} = \frac{\partial\Omega}{\partial x_1} + \frac{\partial\Omega}{\partial x_2} \qquad (5.2)$$

Therefore, equation 3.3 is fulfilled if and only if

$$\frac{\partial\Omega}{\partial x_1}(x_1 , x_2) = 0 \qquad (5.3.a)$$

$$\frac{\partial\Omega}{\partial x_2}(x_1 , x_2) = 0. \qquad (5.3.b)$$

Here the direct sum of D_1 and D_2 has been identified with their cartesian product, in view of the fact that D_1, D_2 constitute a decomposition of D. Accordingly, every $x \varepsilon D$ has here been written as $x = (x_1, x_2)$ where $(x_1, 0) \varepsilon D_1$ and $(0, x_2) \varepsilon D_2$.

Using this notation a dual variational principle for linear operators can be formulated.

5.1 *Theorem*

Let $L : D \to D^*$ be a linear potential operator and Ω be given by equation 3.6. Assume further that a decomposition D_1, D_2 of D exists which satisfies relations 5.1. Define the sets

$$A = \{x_a = (x_{1a} , x_{2a}) \mid x_a \text{ satisfies equation 5.3.a}\} \qquad (5.4.a)$$

and

$$B = \{x_b = (x_{1b} , x_{2b}) \mid x_b \text{ satisfies equation 5.3.b)}\}. \qquad (5.4.b)$$

Then:

i) Each $x_a \varepsilon A$ and $x_b \varepsilon B$ satisfy

$$\Omega(x_{1a} , x_{2a}) \leq \Omega(x_{1b} , x_{2b}). \qquad (5.5)$$

ii) When a solution $\xi = (\xi_1, \xi_2) \varepsilon D$ exists which satisfies the system of equation 5.3 (or equivalently equation 3.1), we have:

α) The maximum in the set A of Ω is achieved at ξ

β) The minimum in the set B of Ω is achieved at ξ

γ) The maximum value of Ω in A coincides with its minimum value in B.

Proof.- First, a convenient notation will be introduced. Given any two elements $x_+ = (x_{1+} , x_{2+}) \varepsilon D$ and $x_- = (x_{1-} , x_{2-}) \varepsilon D$, define:

$$\Delta x_1 = x_{1+} - x_{1-} \qquad (5.6.a)$$

$$\Delta x_2 = x_{2+} - x_{2-} \qquad (5.6.b)$$

$$\Delta\Omega = \Omega(x_{1+} , x_{2+}) - \Omega(x_{1-} , x_{2-}) \qquad (5.6.c)$$

$$\frac{\partial\Omega}{\partial x_{1-}} = \frac{\partial\Omega}{\partial x_1}(x_{1-}, x_{2-}) \tag{5.6.d}$$

$$\frac{\partial\Omega}{\partial x_{2+}} = \frac{\partial\Omega}{\partial x_2}(x_{1+}, x_{2+}) \tag{5.6.e}$$

Now let Ψ be the functional given by equation 3.5. Then

$$\psi(\Delta x_1) - \psi(\Delta x_2) = \Delta\Omega - \langle\frac{\partial\Omega}{\partial x_{1-}}, \Delta x_1\rangle - \langle\frac{\partial\Omega}{\partial x_{2+}}, \Delta x_2\rangle \geq 0 \tag{5.7}$$

Thus, if $x_- \varepsilon A$ and $x_+ \varepsilon B$ it follows that

$$\Omega(x_+) \geq \Omega(x_-) \tag{5.8}$$

and consequently equation 5.5.

On the other hand, if a solution ξ of the system equation 5.3 exists, then necessarily $\xi \varepsilon A \cap B$. In view of this fact, α and β are imme diate consequences of equation 5.5. Finally, γ is now obvious.

5.2 *Corollary*

If in Theorem 5.1, L is assumed to be strictly positive in D_1 and strictly negative in D_2, then

$$i) \quad \Omega(x_a) < \Omega(x_b) \tag{5.9}$$

whenever $x_a \varepsilon A$, $x_b \varepsilon B$ and $x_a \neq x_b$

ii) There is at most one solution of the system equations 5.3.

iii) If the maximum value on A and the minimum value on B of Ω, are both achieved and they are equal, then the solution of the system equation 5.3 exists.

Proof.- The proof of *i)* is straight forward, because in equation 5.7 only the strict inequality can hold when $\Delta x_1 \neq 0$ or $\Delta x_2 \neq 0$. *ii)* follows from *i)* because any solution belongs to A and B. Finally, if the maximum on A is achieved on x_a and the minimum on B is achieved on x_b, and

$$\Omega(x_a) = \Omega(x_b)$$

then x_a and x_b cannot be different by *i)*. Consequently $x_a = x_b \varepsilon A \cap B$ and this is a solution of the system equations 5.3.

The proof of Theorem 5.1, was based on relation equation 5.7. Consequently the generalization of the results of that Theorem to non-linear operators can be carried out for functionals satisfying that relation. The notation introduced in the proof of Theorem 5.1, will be retained in what follows.

5.3 *Definition*

Let $D = D_1 \times D_2$ be the product space of two linear spaces D_1 and D_2. Consider a functional $X : D \to R^1$ which possess an additive Gateaux variation for every $x \varepsilon D$. Then X is said to be saddle, convex in D_1 and concave in D_2, if

$$\Delta X - \langle \frac{\partial X}{\partial x_{1-}}, \Delta x_1 \rangle - \langle \frac{\partial X}{\partial x_{2+}}, \Delta x_2 \rangle \geq 0 \qquad (5.10)$$

for every x_+ , $x_- \varepsilon D$. X is said to be strictly saddle if the inequality is strict whenever $x_+ \neq x_-$.

There is a close connection between the concepts of saddle functional and those of convex and concave functionals.

5.4 *Theorem*

A necessary and sufficient condition for $X : D \to R^1$ to be (strictly) saddle, convex in D_1 and concave in D_2 is that for each fixed $x_2 \varepsilon D_2$, X be (strictly) convex in D_1 and for each fixed $x_1 \varepsilon D_1$, X be (strictly) concave in D_2.

Proof.- Assume X is saddle. Then if $x_{2-} = x_{2+}$, relation equation 5.10 becomes:

$$X(x_{1+} , x_{2-}) - X(x_{1-} x_{2-}) - <\frac{\partial X}{\partial x_{1-}} , \Delta x_1> \geq 0 \qquad (5.11.a)$$

Similarly, if $x_{1-} = x_{1+}$, then:

$$X(x_{1+} , x_{2+}) - X(x_{1+} , x_{2-}) - <\frac{\partial X}{\partial x_{2+}} , \Delta x_2> \geq 0 \qquad (5.11.b)$$

The inequality equation 5.11.a shows that X is convex in x_1 while equation 5.11.b shows in view of equation 4.11 that X is concave in x_2. This completes the proof of necessity. Conversely, if the inequalities equations 5.11 hold simultaneously, by adding them one obtains equation 5.10. All these arguments can be carried out when only the strict inequalities hold and therefore the assertions of the theorem are valid when the strict properties are conside_ red.

Let $X : D \to R^1$ possess an additive Gateaux variation in D. The condition that the Gateaux variation of X vanishes at a point $x = (x_1, x_2) \varepsilon D = D_1 \times D_2$ is equivalent to the system:

$$\frac{\partial X}{\partial x_1}(x_1, x_2) = 0 \qquad (5.12.a)$$

$$\frac{\partial X}{\partial x_2}(x_1, x_2) = 0 \qquad (5.12.b)$$

For such a functional the following theorem is established.

5.5 *Theorem*

Let D_1, D_2 be two linear spaces and $D = D_1 \times D_2$. Consider a functional $X : D \to R^1$ which has a Gateaux variation in D. Assume X is saddle, convex in D_1 and concave in D_2. Define the sets:

$$A = \{x_a = (x_{1a}, x_{2a}) \varepsilon D \mid x_a \text{ satisfies equation 5.12.a}\} \qquad (5.13.a)$$

and

$$B = \{x_b = (x_{1b}, x_{2b}) \varepsilon D \mid x_b \text{ satisfies equation 5.12.b}\} \qquad (5.13.b)$$

Then:

i) Each $x_a \varepsilon A$ and $x_b \varepsilon B$ satisfy

$$\Omega(x_a) \leq \Omega(x_b) \qquad (5.14)$$

ii) When a solution $\xi = (\xi_1, \xi_2) \varepsilon D$ of the system equations 5.12 exists, we have:

α) The maximum in the set A of X is achieved at ξ

ν) The minimum in the set B of X is achieved at ξ

γ) The maximum value of X in A coincides with its minimum value in B.

Proof.- This can be carried out along the same lines as the proof of Theorem 5.1.

5.6 *Corollary*

If in Theorem 5.1, X is assumed to be strictly saddle, then:

i)

$$X(x_a) < X(x_b) \tag{5.15}$$

whenever $x_a \varepsilon A$, $x_b \varepsilon B$ and $x_a \neq x_b$;

ii) There is at most one solution of the system equations 5.12

iii) If the maximum value on A and the minimum value on B of X, are both achieved and are equal, then the solution of the system equations 5.12 exists.

Proof.- This is the same as that of Corollary 5.2.

Theorem 5.5 is a natural generalization of Theorem 5.1. As a matter of fact, the requirement that the functional X be saddle, is nothing else but the requirement that a decomposition of D into D_+ and D_- hold locally, as can be seen in the Theorem that follows.

5.7 *Theorem*

Let $X : D_1 \times D_2 \to R^1$ be such that

i) X has additive Gateaux variations up to the second order

ii) The operator $X'' : D \to D^{2*}$ is bidimensionally continuous in D.

Then a necessary and sufficient condition for X to be (strictly) saddle, convex with respect to D_1 and concave with respect to D_2, is that for every $x \varepsilon D$, $\frac{\partial^2 X}{\partial x_1^2}(x) \varepsilon D^{2*}$ be (strictly) positive and that at the same time for every $x \varepsilon D$, $\frac{\partial^2 X}{\partial x_2^2}(x) \varepsilon D^{2*}$ be (strictly) negative.

Proof.- This follows from Theorem 5.4 in view of Theorem 4.7.

6. HAMILTON PRINCIPLES

In this section Hamilton principles {6 - 11} will be derived from the results of section 5.

Let $D = D_1 \times D_2$ where D_1 and D_2 are two linear spaces. Consider an operator $\hat{T} : D \to D^*$ and define a functional $\theta : D \to R^1$ given for every $x = (x_1, x_2) \varepsilon D$ by:

$$\theta(x) = \langle \hat{T} x_1 , x_2 \rangle = \langle \hat{T}^* x_2 , x_1 \rangle \tag{6.1}$$

Here $\hat{T}^*$ is the adjoint of $\hat{T}$. Given a functional $X : D \to R^1$ another functional $\hat{X} : D \to R^1$ is defined, given for every $x = (x_1, x_2) \varepsilon D$ by:

$$\hat{X}(x) = X(x) - \theta(x) \tag{6.2}$$

When Theorem 5.5 and Corollary 5.6 are applied to the functional $\hat{X}$, one gets Hamilton's dual variational principles as given by Noble and Sewell {11}. This can be seen in the next theorem.

6.1 Theorem

Let X possess an additive Gateaux variation on D and $\hat{X}$ be given as explained before. Then there exists a linear operator $T : D \to D^*$ such that:

$$\frac{\partial \hat{X}}{\partial x_1}(x_1, x_2) = \frac{\partial X}{\partial x_1}(x_1, x_2) - T^* x_2 \qquad (6.3.a)$$

$$\frac{\partial X}{\partial x_2}(x_1, x_2) = \frac{\partial X}{\partial x_2}(x_1, x_2) - T x_1 \qquad (6.3.b)$$

In addition, any $x_1 \varepsilon D_1$ and $x_2 \varepsilon D_2$ satisfies:

$$\langle Tx_1, x_1 \rangle = 0 \qquad (6.4.a)$$

$$Tx_2 = 0 \qquad (6.4.b)$$

$$T^* x_1 = 0 \qquad (6.4.c)$$

$$\langle T^* x_2, x_2 \rangle = 0 \qquad (6.4.d)$$

Proof.- Before proceeding to prove this theorem, observe that equations 6.4 show that T and T* can be interpreted as mappings $T : D_1 \to D_2^*$ and $T^* : D_2 \to D_1^*$ such that:

$$\langle Tx_1, x_2 \rangle = \langle T^* x_2, x_1 \rangle \qquad (6.5)$$

for any $x_1 \varepsilon D_1$ and $x_2 \varepsilon D_2$. This interpretation yields Rall's formulation {17}.

To prove the theorem, define $T : D \to D^*$ for every $x = (x_1, x_2) \varepsilon D$ and $y = (y_1, y_2) \varepsilon D$ by:

$$<T(x), y> = <\hat{T}x_1 , y_2> \qquad (6.6.a)$$

Consequently T is linear and

$$<T^*x, y> = <\hat{T}y_1 , x_2> \qquad (6.6.b)$$

Using these definitions of T and T^* the proof of the theorem is straight forward.

Finally, it is worthwhile recalling that the correspondence $X \leftrightarrow \hat{X}$ defined by equation 6.2, which is obviously one to one, preserves the saddle property, because:

$$\Delta X - <\frac{\partial X}{\partial x_{1-}} , \Delta x_1> - <\frac{\partial X}{\partial x_{2+}} , \Delta x_2> =$$

$$\Delta \hat{X} - <\frac{\partial \hat{X}}{\partial x_{1-}} , \Delta x_1> - <\frac{\partial \hat{X}}{\partial x_{2+}} , \Delta x_2> \qquad (6.7)$$

7. APPLICATIONS

Two applications will be made of the foregoing theory; in the first boundary conditions of linear problems will be treated and in the second dual variational principles for initial value problems of the heat equation will be developed.

i) Boundary conditions for linear problems. Boundary conditions that have been treated by previous authors {11, 18, 19} can be included within the present framework and a general formulation for them will be presented elsewhere. Here, only the general ideas of the procedure are explained and applied to a linear problem.

Let D be a linear space of functions defined on a compact connected subset E of R^n with smooth boundary ∂E. Let $\hat{L} : D \to D^*$ be a linear operator. A formal adjoint of $\hat{L}$ is an operator $\hat{L}^{\#} : D \to D^*$ such that for any x, $y \varepsilon D$:

$$<\hat{L}x\,,\, y> \;=\; <\hat{L}^{\#}y\,,\, x> + \;\; \phi(x\,,\, y) \qquad (7.1)$$

where $\phi : D^2 \to R^1$ is a functional whose support is contained in ∂E. Obviously, ϕ is a bilinear functional in D, because $<\hat{L}x\,,\,y>$ and $<\hat{L}^{\#}y\,,\,x>$ are bilinear functionals. It is therefore possible to define operators $\hat{B} : D \to D^*$ and $\hat{B}^{\#} : D \to D^*$ with support in ∂E such that for every x, $y \varepsilon D$:

$$<\hat{L}x\,,y> + <\hat{B}x\,,y> \;=\; <\hat{L}^*y\,,x> + <\hat{B}^{\#}y\,,x> \qquad (7.2)$$

Indeed, all that is required is that

$$<\hat{B}^{\#}y\,,x> - <\hat{B}x\,,y> \;=\; \phi(x,\; y) \qquad (7.3)$$

for every x, $y \varepsilon D$. In view of equation 7.2 an operator $L : D \to D^*$ and its adjoint $L^* : D \to D^*$ can be defined by:

$$<Lx\,,y> \;=\; <\hat{L}x\,,y> + <\hat{B}x\,,y> \qquad (7.4.a)$$

and

$$\langle L^{*}x,\ y\rangle = \langle \hat{L}^{\#}x,\ y\rangle + \langle \hat{B}^{\#}x,\ y\rangle \qquad (7.4.b)$$

which hold for every $x,\ y \in D$.

In applications, the equation

$$\hat{L}x = \hat{f} \qquad (7.5)$$

can frequently be established at the outset, together with some linear boundary conditions. To formulate the problem in terms of functional valued operators it is necessary to construct $L : D \to D^{*}$ and $f \in D^{*}$ in such a way that

$$Lx = f \qquad (7.6)$$

if and only if equation 7.5 and the boundary data are satisfied. Assume that $\hat{L}x + \hat{B}x = 0$ if and only if $\hat{L}x = 0$ and $\hat{B}x = 0$. Then a further requirement for $\hat{B}$ is that $\hat{B}x$ vanishes if and only if x satisfies zero boundary conditions. In this case

$$f = \hat{f} + \hat{f}_B \qquad (7.7)$$

where $f_B \in D^{*}$ is determined by the boundary data alone.

However, it can be seen that even after satisfying these requirements, considerable arbitriness is left for the definition of the adjoint operator. For example, if L is self-adjoint, then equation 7.3 becomes:

$$\langle \hat{B}y,\ x\rangle - \langle \hat{B}x,\ y\rangle = \phi(x,\ y) \qquad (7.8)$$

which obviously implies that $\phi(x, y)$ is anti-symmetric. Let $P : D \to D^*$ be self-adjoint, such that $Px = 0$ whenever x satisfies zero boundary conditions. Then

$$L + P = \hat{L} + \hat{B} + P \tag{7.9}$$

is also self-adjoint and can also be used to define the functional equation which is equivalent to the given problem.

To illustrate the preceding abstract theory, consider the heat equation

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = f_R(x, t) \tag{7.10.a}$$

in a normalized region $0 \leq t \leq 1$ and $0 \leq x \leq 1$, with boundary condition

$$u(0, t) = f_{B0};\ 0 \leq t \leq 1 \tag{7.10.b}$$

$$u(1, t) = f_{B1};\ 0 \leq t \leq 1 \tag{7.10.c}$$

and initial condition

$$u(x, 0) = u_o(x)\ ;\ 0 \leq x \leq 1 \tag{7.10.d}$$

The linear space in which the problem is set, will be taken to be

$$D = \{u: [0,1] \times [0,1] \to R^2 | u \varepsilon C^2\} \tag{7.11}$$

It has been shown previously {5} that in order to associate to this problem a symmetric functional valued operator L, it is convenient to introduce a transformation p which reflects the variable t of any function at the mid-point of the interval [0,1]; i.e. $(pw)(t) = w(1-t)$

Then applying p to equation 7.10.a the equivalent equation

$$p\,\frac{\partial u}{\partial t} - p\,\frac{\partial^2 u}{\partial x^2} = pf_R \tag{7.12}$$

is obtained.

Integration of the left-hand member of equation 7.12 after multiplying by $v \varepsilon D$ leads to define $\hat{L} : D \rightarrow D^*$ by

$$<\hat{L}u,\ v> = \int_0^1 \left[v * \frac{\partial u}{\partial t} - v * \frac{\partial^2 u}{\partial x^2} \right] dx \tag{7.13}$$

which holds for every $u,\ v \varepsilon D$. Here the notation

$$(u * v)\ (x) = \int_0^1 u(x,t)\ v(x,1-t)dt \tag{7.14}$$

has been used. Integration by parts in equation 7.13 yields

$$<\hat{L}u,\ v> - <\hat{L}v,\ u> = \int_0^1 \left[u(x,\ 1)\ v(x,\ 0) - u(x,\ 0)\ v(x,\ 1) \right] dx$$

$$+ \left[\frac{\partial u}{\partial x} * v - \frac{\partial v}{\partial x} * u \right]_{x=0} - \left[\frac{\partial u}{\partial x} * v - \frac{\partial v}{\partial x} * u \right]_{x=1} \tag{7.15}$$

Clear this equation leaving on only one side all the data for u of the problem, to obtain:

$$\langle \hat{L}u, v\rangle + \left[u * \frac{\partial v}{\partial x}\right]_{x = 0} - \left[u * \frac{\partial v}{\partial x}\right]_{x = 1} + \int_0^1 u(x, 0)\ v(x, 1)dx$$

$$= \langle \hat{L}v, u\rangle + \int_0^1 \left[v(x, 0)\ u(x, 1)\right] dx + \left[v * \frac{\partial u}{\partial x}\right] - \left[v * \frac{\partial u}{\partial x}\right]_{x = 1} \qquad (7.16)$$

The left-hand side of this equation defines a functional of v. It is the null functional if and only if u is a solution of problem 7.10 with zero data.

Therefore $\hat{B}$ can be defined by

$$\langle \hat{B}u, v\rangle = \left[u * \frac{\partial v}{\partial x}\right]_{x = 0} - \left[u * \frac{\partial v}{\partial x}\right]_{x = 1} + \int_0^1 u(x, 0)\ v(x, 1)dx \qquad (7.17)$$

and $L = \hat{L} + \hat{B}$. Applying equations 7.13 and 7.17 to a solution of the system 7.10 leads to define $\hat{f}$ and $\hat{f}_B$ by

$$\langle \hat{f}, v\rangle = \int_0^1 f_R * v\ dx \qquad (7.18.a)$$

and

$$\langle \hat{f}_B, v\rangle = u_{B0} * \left(\frac{\partial v}{\partial x}\right)_{x = 0} - u_{B1} * \left(\frac{\partial v}{\partial x}\right)_{x = 1} + \int_0^1 u_0(x)\ v(x, 1)dx \qquad (7.18.b)$$

for every $v \varepsilon D$. In this manner equation 7.6 is fulfilled if and only if equations 7.10 are satisfied. By inspection of equation 7.16 it is seen that L is self-adjoint. However, let $P : D \rightarrow D^*$ be given by

$$\langle Pu, v\rangle = \int_0^1 u(x, 0)\ v(x, 0)dx \qquad (7.19)$$

and $f_p \varepsilon D^*$ by

$$\langle f_p, v\rangle = \int_0^1 u_0(x)\ v(x,\ 0)dx \tag{7.20}$$

Then $L + P$ is self-adjoint and

$$(L + P)u = f + f_p \tag{7.21}$$

is equivalent to equation 7.10. In a similar manner additional terms associated with the boundary conditions can be introduced.

Define now

$$X(u) = \frac{1}{2} \langle Lu,\ u\rangle - \langle f,\ u\rangle \tag{7.22}$$

Theorem 3.1 implies that u satisfies the system 7.10 if and only if

$$X'(u) = 0 \tag{7.23}$$

This result was presented in {5} but now dual variational principles can be formulated.

ii) Dual variational principles. For simplicity take D as before, except that now its elements are required to satisfy homogeneous boundary conditions (equations 7.10.b and c). Consequently in what follows f_{B0} and f_{B1} will be assumed to be indentically zero. In this case

$$X(u) = \frac{1}{2} \int_0^1 \left[u * \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} * \frac{\partial u}{\partial x} - 2f_R * u\right]dx$$

$$+ \frac{1}{2} \int_0^1 \left[u(x,\ 0) - 2u_0(x)\right] u(x,1)dx \tag{7.24}$$

The sets

$$D_1 = \{u \varepsilon D \mid u(1 - t) = u(t)\} \tag{7.25.a}$$

and

$$D_2 = \{u \varepsilon D \mid u(1 - t) = - u(t)\} \tag{7.25.b}$$

constitute a decomposition of D. In addition for every $u \varepsilon D_1$

$$\langle Lu, u\rangle = \frac{1}{2} \int_0^1 \left\{ \int_0^1 \left[\frac{\partial u}{\partial x}\right]^2 dt + u^2(x, 0)\right\} dx \leq 0 \tag{7.26.a}$$

and for every $u \varepsilon D_2$

$$\langle Lu, u\rangle = - \frac{1}{2} \int_0^1 \left\{ \int_0^1 \left[\frac{\partial u}{\partial x}\right]^2 dt + u^2(x, 0)\right\} dx \leq 0 \tag{7.26.b}$$

Furthermore, in equation 7.26 the equalities hold only if u is identically zero. Thus, L is strictly positive in D_1 and strictly negative in D_2. Consequently Theorem 5.1 and Corollary 5.2 are applicable.

To obtain the dual equations, observe that

$$\langle X'(u), v\rangle = \langle Lu - f, v\rangle \tag{7.27}$$

Then

$$\langle \frac{\partial X}{\partial u_1} (u), v\rangle = \langle Lu - f, v_1\rangle \tag{7.28.a}$$

and

$$<\frac{\partial X}{\partial u_2}(u),\ v> \ = \ < Lu - f,\ v_2> \qquad (7.28.b)$$

Consequently equations 5.3 become

$$\left.\begin{aligned} \left[\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2}\right]_S &= \left[f_R\right]_S \\ u(x,\ 0) &= u_0(x) \end{aligned}\right\} \qquad (7.29.a)$$

$$\left.\begin{aligned} \left[\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2}\right]_A &= \left[f_R\right]_A \\ u(x,\ 0) &= u_o(x) \end{aligned}\right\} \qquad (7.29.b)$$

APPENDIX

In this appendix the proof of Theorem 2.5 is given. Before proceeding to develop it, a few remarks will be made.

For any fixed x, $y \varepsilon D$, observe that $\langle P(x + t(y - x)), y - x \rangle$ is a continuous function of the real variable t. Consequently, this function can be integrated with respect to t and a functional $G : D \times D \rightarrow R^1$ can be defined by

$$G(x, y) = \int_0^1 \langle P(x + t(y - x)), y - x \rangle dt \qquad (A.1)$$

A.1 LEMMA

Under the hypotheses of Theorem 2.5, the functional G given by equation A.1 satisfies the relation

$$G(x, z) = G(x, y) + G(y, z) \qquad (A.2)$$

for every x, y, $z \varepsilon D$, when $P'(x) \varepsilon D^{2*}$ is symmetric for every $x \varepsilon D$.

Proof.- Consider two functions of the real variables ξ, η defined by

$$F_1(\xi, \eta) = \langle P(x + \xi(y - x) + \eta(z - y)), y - x \rangle \qquad (A.3.a)$$

$$F_2(\xi, \eta) = \langle P(x + \xi(y - x) + \eta(z - y)), z - y\rangle \qquad \text{(A.3.b)}$$

Obviously

$$\frac{\partial F_1}{\partial \eta}(\xi, \eta) = \frac{\partial F_2}{\partial \xi}(\xi, \eta) \qquad \text{(A.4)}$$

because P' is symmetric. Consequently the line integral $\int F_1 d\xi + F_2 d\eta$ vanishes for any closed curve on the ξ, η-plane. But equation A.2 is now obvious when this relation is applied to the triangle joining the points $(0, 0)$, $(1,0)$ and $(1,1)$ of the ξ, η-plane.

Proof of Theorem 2.5.- Let $\Psi : D \to R^1$ be a potential of P, and given any $x, y, z \varepsilon D$ define a function of two real variables by

$$f(\xi, \eta) = \Psi(x + \xi y + \eta z) \qquad \text{(A.5)}$$

Under the hypotheses of the theorem, $f \varepsilon C^2$ and consequently $\frac{\partial^2 f}{\partial \xi \partial \eta} = \frac{\partial^2 f}{\partial \eta \partial \xi}$. Using this fact it may easily be seen that the condition is necessary.

To show that the condition is sufficient, the functional G given by equation A.1 will be used. In addition, it is convenient to introduce a functional $\Psi : D \to R^1$ given by

$$\Psi(x) = \int_0^1 \langle P(tx), x\rangle \, dt = G(0, x) \qquad \text{(A.6)}$$

for every $x \in D$. If given $x, y \in D$, a function g of the real variable t is defined by

$$g(t) = G(x, x + ty) = t\int_0^1 <P(x, t\tau y), y> d\tau \qquad (A.7)$$

then

$$g'(0) = <P(x), y> \qquad (A.8)$$

because the expression under the integral sign in equation A.7 possesses a first order continuous derivative.

Finally, given $x, y \in D$, let $h : R^1 \to R^1$ be given by

$$h(t) = \Psi(x + ty) = G(0, x) + G(x, x + ty) \qquad (A.9)$$

for every real t. Then

$$h'(0) = <P(x), y> \qquad (A.10)$$

which shows that the Gateaux variation of Ψ exists, is additive and

$$\Psi'(x) = P(x) \qquad (A.11)$$

REFERENCES

1. M. Z. Nashed, *Differentiability and related properties of nonlinear operators: some aspects of the role of differentials in nonlinear functional analysis*, In L. B. Rall (Ed.), Nonlinear functional analysis and applications, pp 103-309, Academic Press, New York (1971)

2. M. M. Vainberg, *Variational methods for the study of nonlinear operator equations*, Holden Day, San Francisco (1964)

3. E. Tonti, *A systematic approach to the serarch for variational principles*, International conference on variational methods in engineering, Southampton University, England (1972), pp 1.1-1.12

4. E. Tonti, *On the variational formulation for linear initial value problems*, Annali di Matematica pura ed applicata, Vol 95 (1973), pp 331-360

5. I. Herrera and J. Bielak, *A simplified version of Gurtin's variational principles*, Arch. Rational Mech. Anal., Vol 53 (1974), pp 131-149

6. B. Noble, *Complementary variational principles for boundary value problems I. Basic principles*, Report No 473, Mathematics Research Center, University of Wisconsin (1964)

7. B. Noble, *Complementary variational principles II. Nonlinear networks*, Report No 643, Mathematics Research Center, University of Wisconsin (1966)

8. M. J. Sewell, *On dual approximation principles and optimization in continuum mechanics*, Phil. Trans. Royal Society, London, Vol 265 (1969), pp 319-351

9. A. M. Arthurs, *Complementary variational principles*, Oxford University Press (1970)

10. P. D. Robinson, *Complementary variational principles*, In L. B. Rall (Ed.), Nonlinear functional analysis and applications, Academic Press, New York (1971), pp 507-576

11. B. Noble and M. J. Sewell, *On dual extremum principles in Applied Mathematics*, Journal Inst. Mathematics Applics., Vol 9 (1972), pp 123-193

12. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol 1, Interscience, New York (1953)

13. J. L. Synge, *The hypercicle in Mathematical Physics*, Cambridge University Press (1957)

14. J. B. Diaz, *Upper and lower bounds for quadratic integrals, and at a point, for solutions of linear boundary value problems*. Procs., Symposium on Boundary Value Problems in Differential Equations, Mathematics Research Center, University of Wisconsin Press (1959)

15. T. Kato, *On some approximate methods concerning the operators* T*T, Math. Ann. 126 (1953), pp 253-262

16. H. Fujita, *Contribution to the theory of upper and lower bounds in boundary value problems*, Journal Physics Society, Vol 10, No 1, Japan (1955)

17. L. B. Rall, *On complementary variational principles*, Journal Mathematics Analysis Applics, Vol 14 (1966), pp 174-184

18. V. Komkov, *Application of Rall's theorem to classical elastodynamics*, Journal Mathematics Analysis Applics, Vol 14 (1966), pp 511-521

19. A. M. Arthurs, *A note on Komkov's class of boundary value problems and associated variational principles*, Journal Mathematics Analysis Aplics, Vol 33 (1971), pp 402-407