56

J. Inst. Maths Applics (1978) **21**, 95–115

Dual Extremum Principles for Non-negative Unsymmetric Operators

I. HERRERA

National University of Mexico

AND

M. J. SEWELL

Department of Mathematics, University of Reading

[Received 21 April 1977]

Dual extremum principles are constructed for non-negative unsymmetric operator equations. The theory is in terms of functional-valued linear operators defined on an infinite-dimensional space which need not even be normed. Saddle operators and saddle functionals are constructed as essential ingredients. An extension to affine sub-spaces required for certain partial differential equations is included. New dual extremum principles for versions of the heat equation and the wave equation are stated in illustration.

1. Introduction

MUCH PROGRESS has been made in recent years in the development of a theoretical framework for the formulation of variational principles (Sewell, 1969; Noble & Sewell, 1972; Sewell, 1973*a*,*b*; Herrera, 1974; Herrera & Bielak, 1976; Sewell & Noble, 1976). A contribution due to Sewell (1973*a*) allows one to generate dual variational principles from a single functional and to unify the theory. Herrera (1974), and Herrera & Bielak (1976), developed a formulation of variational principles using functional-valued operators which simplifies the treatment of partial differential equations because it is applicable in any linear space, not necessarily normed, nor with an inner product, nor complete. The situation has been summarized by Herrera & Bielak (1976).

A sufficient condition for an operator equation to admit a variational formulation is that the operator be potential, and this will be so if and only if its derivative is a symmetric bilinear functional. Variational and extremum principles are interrelated because a necessary condition for the existence of a stationary extremum is the vanishing of the variation. A sufficient condition for a variational principle to be extremum is that the generating functional be either convex or concave. In general, a functional is neither convex nor concave; however, if the linear space in which the functional is defined can be decomposed into two subspaces, one in which the functional is convex and another in which it is concave, the functional is saddle, so that Noble & Sewell's results can be applied; then the variational principle becomes a pair of dual principles. Such a decomposition can be expected to exist under very general conditions, at least locally, because this is the case for surfaces in finite dimensional spaces.

For linear operators the theory becomes especially simple. A sufficient condition for a linear equation to admit a variational formulation is that the operator be symmetric. If the operator is either positive or negative, then the associated functional is either convex or concave, and consequently the variational principle is an extremum one. In general a symmetric linear operator is neither positive nor negative, but under very general conditions (e.g. when spectral theorems are applicable), the linear space D in which the operator is defined can be decomposed into two subspaces, one D_+ in which the operator is non-negative and another D_- in which it is non-positive; then the associated functional is saddle and the variational principle generates dual extremum principles of Sewell's type. Accordingly, it is possible to associate dual extremum principles to arbitrary equations formulated in terms of linear symmetric operators.

The main advantages of using functional-valued operators are (Herrera & Bielak, 1976):

(i) Problems are formulated in the most general kind of linear spaces, which are not necessarily normed, nor with an inner product, nor complete. Most work in this field has been done in either inner product spaces (Noble & Sewell, 1972; Sewell, 1969, 1973*a*,*b*; Sewell & Noble, 1976) or in Hilbert spaces (Arthurs, 1970; Robinson, 1971; Collins, 1976; Brezis & Ekeland, 1976) and it is generally thought that this is desirable, if not essential, for the results to hold. In many applications the introduction of the Hilbert space structure leads to unwarranted complications. This is not required when functional-valued operators are used.

(ii) The introduction of superfluous hypotheses in the development of a theory is always inconvenient, because frequently they needlessly restrict its applicability.

(iii) The symmetry condition for the potentialness of an operator can be extended to linear spaces for which no inner product or norm have to be defined (Herrera, 1974). This fact makes it possible to formulate a theory which is rigorous and at the same time not complicated.

(iv) Error bounds for approximate solutions are among the most important results that the theory yields. They depend, however, on simple properties which are independent of any Hilbert space structure, and therefore can be obtained in the simple setting developed by Herrera. General results for such bounds are given in Section 3 of this paper; they induce a metric in the space. By using these bounds it is possible to define weak or generalized solutions; however, these will not be discussed here. Certain error bounds for non-linear problems have been described by Sewell & Noble (1976).

The application of the theory of functional-valued operators to partial differential equations requires one to express them in terms of such operators. This is achieved by repeated use of integration by parts formulae. In this respect, the main advantage over standard formulations is derived from the fact that such manipulations can be carried out more systematically; for example, the operators involved always possess an adjoint. A discussion of the treatment of boundary conditions has already been given (Herrera, 1974; Herrera & Bielak, 1976); and in another way by Noble & Sewell (1972 *et seq.*).

Due to its generality, the usefulness of the concept of saddle functional is great. In applications the flexibility of this concept is enhanced by the fact that its definition (Sewell, 1969, Section 2 (vii)) depends essentially on the system of coordinates used.

This is a fundamental difference with respect to the same concept when applied to surfaces; the saddle property of a surface is independent of the system of coordinates used.

In some applications it may be difficult to find a decomposition of the space with respect to which the functional is saddle. For linear problems (Herrera, 1974), this amounts to constructing a decomposition (D_+, D_-) of the space, such that the operator is non-negative on D_+ and non-positive on D_-. However, the restrictions imposed by these conditions are not too severe, because such a decomposition is not unique, and in this paper a wide class of variational principles is given for which it is easily obtained.

For initial value problems only stationary and not extremum principles had been obtained, until recently. However Herrera (1974) obtained dual extremum principles for initial value problems that were later applied to a large class of problems by Herrera & Bielak (1976). Later Collins (1976) extended these results to a general kind of dissipative system by considering simultaneously the adjoint equation. In France, extremum principles for initial value problems corresponding to certain parabolic equations have just been obtained (Brezis & Ekeland, 1976); they impose more severe restrictions on the boundary conditions than those mentioned above.

The procedure used by Collins for dissipative systems (essentially, the mirror method) to generalize Herrera's principles offers interesting possibilities, because this approach is not limited to dissipative systems. In this paper a general procedure for deriving dual extremum principles by the mirror method is developed; it is applicable to any linear problem formulated in terms of a non-negative (or alternatively non-positive) functional-valued operator. To illustrate the method, it is applied to the heat and the wave equations. To our knowledge, these are the first extremum principles for the initial value problem associated with the wave equation, even though further work will be required to apply them. The principles derived for the heat equation are different from those available up to now and impose less restrictive conditions on the test functions.

A brief description of the theory of functional-valued operators is given in Section 2 and saddle operators are introduced there. Variational and extremum principles for arbitrary saddle operators are given in Section 3 and in the next Section these results are used to develop a general theory of non-negative unsymmetric operators. Section 5 is devoted to extending this theory to problems defined in affine subspaces. Some applications of the theory to partial differential equations are given in Section 6. Only the heat and wave equations are considered, but the application of the theory to other equations can be made in a similar manner.

2. Saddle Operators

The notion of saddle functional (Sewell, 1969; Noble & Sewell, 1972) has been used extensively in the formulation of dual extremum principles and Herrera (1974) has used it to develop an approach adequate for the treatment of partial differential equations (Herrera & Bielak, 1974, 1976). In this section a version of this method which is suitable for linear problems will be described.

Let D be any linear space, not necessarily metric, nor normed, nor with an inner product, nor complete. The value of any real linear functional $f\colon D \to R$ at an element

$v \in D$ is a real number to be denoted by $\langle f, v \rangle$. It is emphasized that this is *not* an inner product, even though the same notation is often used for such (in particular by Noble & Sewell, 1972). The linear space of all such functionals will be denoted by D^*.

Consider a linear operator $P: D \to D^*$ which is therefore a *functional-valued* linear operator. Thus, for each $u \in D$ there is an element $Pu \in D^*$, and the value of this functional at any $v \in D$ is $\langle Pu, v \rangle$. Given any such operator, the adjoint operator $P^*: D \to D^*$, defined by the condition that

$$\langle P^*u, v \rangle = \langle Pv, u \rangle \tag{2.1}$$

is satisfied $\forall u \in D$ and $\forall v \in D$, always exists. The operator P will be said to be symmetric, or self-adjoint, whenever $P^* = P$. It is important to observe that the definition of adjoint operator given here is simpler than some others frequently used; in particular it does not require the definition of D^{**} nor the condition $D^{**} = D$, as the examples in Section 6 show. This is an advantage of the use of functional-valued operators (Herrera, 1974).

Two subspaces $D_1 \subset D$ and $D_2 \subset D$ are called a decomposition of D if D_1 and D_2 are linearly independent and $D = D_1 + D_2$. In this case, given any $u \in D$ there is a unique couple (u_1, u_2) belonging to D_1 and D_2 respectively, such that $u = u_1 + u_2$; the elements u_1 and u_2 are called projections of u on D_1 and D_2. Projections of linear functionals and functional-valued operators are defined similarly. Given $f \in D^*$ and $P: D \to D^*$, the projections $f_1 \in D^*$ and $f_2 \in D^*$ are such that

$$\langle f_1, u \rangle = \langle f, u_1 \rangle \qquad \langle f_2, u \rangle = \langle f, u_2 \rangle \tag{2.2a}$$

for every $u \in D$, while the projections $P_1: D \to D^*$ and $P_2: D \to D^*$ satisfy the conditions

$$P_1 u = (Pu)_1, \qquad P_2 u = (Pu)_2. \tag{2.2b}$$

From these definitions it follows that

$$f = f_1 + f_2 \tag{2.3a}$$

$$P = P_1 + P_2. \tag{2.3b}$$

Definition 2.1. Let $E \subset D$ be a subspace of D. Then an operator $P: D \to D^*$ is said to be non-negative on E if

$$\langle Pu, u \rangle \geqslant 0 \quad \text{for every } u \in E; \tag{2.4}$$

and positive on E if the equality sign in (2.4) holds only when $u = 0$. When $E = D$ the specification "on E" will be omitted.

If P is non-negative on E the non-negative square root of $\langle Pu, u \rangle$ will be denoted by $\|u\|_P$ whenever $u \in E$. The set

$$N_P = \{u \in E \mid \|u\|_P = 0\} \tag{2.5}$$

may contain non-zero vectors. However, N_P is always a linear subspace of $E \subset D$.

LEMMA 2.2. *Let P be non-negative on E. Then N_P is a linear subspace of $E \subset D$.*

Proof. It is required to prove:

(i) $u \in N_P \Rightarrow \alpha u \in N_P, \quad \forall \alpha \in R$

(ii) $u, v \in N_P \Rightarrow u + v \in N_P$.

Property (i) is obvious because

$$\langle P(\alpha u), \alpha u \rangle = \alpha^2 \langle Pu, u \rangle.$$

Property (ii) follows from

$$\tfrac{1}{2}\langle P(u+v), u+v\rangle + \tfrac{1}{2}\langle P(u-v), u-v\rangle = \langle Pu, u\rangle + \langle Pv, v\rangle \tag{2.6}$$

which can be obtained expanding the left-hand member of this equation. Recalling that each of the terms involved in this equation is non-negative, the property is clear.

Definition 2.3. Let $P: D \to D^*$ be non-negative on E. Then $u, v \in D$ are said to be P-equivalent if and only if $u-v \in N_P$. In this case one writes:

$$u \approx_P v. \tag{2.7}$$

In view of Lemma 2.2, this is an equivalence relation.

Non-positive and negative operators are defined by reversing the sense of the inequality (2.4). For this kind of operator there are definitions and results corresponding to those already given for non-negative and positive operators. It is worth noticing that the relation $\approx_P$ becomes equality when P is either positive or negative on E.

Definition 2.4. Let D_1, D_2 be a decomposition of D and $P: D \to D^*$ be a self-adjoint operator. Then, P is said to be saddle with respect to D_1, D_2 if P is non-negative on D_1 and non-positive on D_2. It is strictly saddle if P is positive on D_1 and negative on D_2.

It must be observed that in this definition the sub-spaces cannot be interchanged. However, this manner of introducing them simplifies many of the propositions to be presented. For saddle operators the bilinear functional $\langle Pu_1, v_1\rangle - \langle Pu_2, v_2\rangle$ is symmetric and non-negative. The non-negative square root of $\langle Pu_1, u_1\rangle - \langle Pu_2, u_2\rangle$ will be denoted by $\|u\|_P$; when P is strictly saddle $\|u\|_P$ is positive and therefore $\|u\|_P$ is a norm. In this case the bilinear functional $\langle Pu_1, v_1\rangle - \langle Pu_2, v_2\rangle$ is an inner product.

Definition 2.5. Let $P: D \to D^*$ be saddle with respect to the decomposition D_1, D_2 of D. Then, $u, v \in D$ are said to be P-equivalent if and only if $\|u-v\|_P = 0$; in this case one writes

$$u \approx_P v.$$

For saddle operators the set

$$N_P = \{u \in D \mid \|u\|_P = 0\}$$

is a subspace of D. Thus relation (2.8) is an equivalence relation.

The following lemma will be used in the construction of dual extremum principles.

LEMMA 2.6. *Let $P: D \to D^*$ be saddle with respect to the decomposition D_1, D_2. Define the subspaces of D*

$$D_a = \{u \in D \mid P_1 u = 0\} \tag{2.9a}$$

$$D_b = \{u \in D \mid P_2 u = 0\}. \tag{2.9b}$$

Then the operator P is non-positive in D_a and non-negative in D_b. Even more, if $u \in D_a$

$$\langle Pu, u\rangle = -\|u\|_P^2. \tag{2.10a}$$

If $u \in D_b$

$$\langle Pu, u\rangle = \|u\|_P^2. \tag{2.10b}$$

Proof. Assume $u \in D_a$, and let u_1 and u_2 be its projections. Then

$$\langle Pu, u_1\rangle = 0$$

and therefore

$$\langle Pu_2, u_1\rangle = \langle Pu, u_1\rangle - \langle Pu_1, u_1\rangle = -\langle Pu_1, u_1\rangle.$$

Hence

$$\langle Pu, u\rangle = \langle Pu, u_1\rangle + \langle Pu_1, u_2\rangle + \langle Pu_2, u_2\rangle$$
$$= \langle Pu_2, u_2\rangle - \langle Pu_1, u_1\rangle = -\|u\|_P \leqslant 0.$$

The proof of the other part of the lemma is similar.

In some applications it is relevant to consider operators which are saddle on a subspace only. Let $E \subset D$ be a subspace, and E_1, E_2 a decomposition of E; then P: $D \to D^*$ is said to be saddle on E with respect to the decomposition E_1, E_2, if P is non-negative on E_1 and non-positive on E_2.

Strictly saddle on E is defined correspondingly.

3. Variational and Extremum Principles

In this section a linear space $\hat{D}$ will be considered whose elements will be represented by $\hat{u}$. Following Herrera (1974), the concept of derivative of a functional will be used in the sense of additive Gateaux variation.

Definition 3.1. Let X: $\hat{D} \to R$ be an arbitrary functional and $\hat{u} \in \hat{D}$. Then $X'(\hat{u}) \in D^*$ is said to be the derivative of X at $\hat{u}$ if

$$\frac{d}{d\lambda} X(\hat{u}+\lambda\hat{v})\,|_{\lambda=0} = \langle X'(\hat{u}), \hat{v}\rangle \tag{3.1}$$

for every $\hat{v} \in \hat{D}$.

The partial derivatives can be defined in terms of the total derivative.

Definition 3.2. Let $\hat{D}_{\mathrm{I}}, \hat{D}_{\mathrm{II}}$ be a decomposition of $\hat{D}$. Assume the derivative of the functional X: $\hat{D} \to R$ exists at $\hat{u} \in \hat{D}$. Then the partial derivatives $X,_{\mathrm{I}}(\hat{u}) \in D^*$ and $X,_{\mathrm{II}}(\hat{u}) \in D^*$ are defined as the projection of $X'(\hat{u}) \in D^*$, i.e. in terms of (2.2)

$$X,_{\mathrm{I}}(\hat{u}) = \{X'(\hat{u})\}_{\mathrm{I}} \quad \text{and} \quad X,_{\mathrm{II}}(\hat{u}) = \{X'(\hat{u})\}_{\mathrm{II}}. \tag{3.2}$$

Associated with a linear functional-valued operator $\hat{P}$: $\hat{D} \to \hat{D}^*$ and for a given $\hat{f} \in \hat{D}^*$, consider the equation

$$\hat{P}\hat{u} = \hat{f}. \tag{3.3}$$

When $\hat{P}$ is self-adjoint it is possible to establish a variational principle. To this end, define the functional X: $\hat{D} \to R$ by

$$X(\hat{u}) = \tfrac{1}{2}\langle \hat{P}\hat{u}, \hat{u}\rangle - \langle \hat{f}, \hat{u}\rangle. \tag{3.4}$$

When X is given by (3.4), the derivative $X'(\hat{u})$ always exists and is

$$X'(\hat{u}) = \hat{P}\hat{u} - \hat{f}. \tag{3.5}$$

In view of this equation, the following theorem is obvious.

THEOREM 3.3. *Let $\hat{P}$: $\hat{D} \to \hat{D}^*$ be self adjoint. Then $\hat{u} \in \hat{D}$ satisfies* (3.3) *if and only if*

$$X'(\hat{u}) = 0. \tag{3.6}$$

However, this is not an extremum principle. When there is available a decomposition $\hat{D}_{\mathrm{I}}, \hat{D}_{\mathrm{II}}$ of $\hat{D}$ such that $\hat{P}$ is saddle with respect to $\hat{D}_{\mathrm{I}}, \hat{D}_{\mathrm{II}}$, it is possible to formulate variational principles that are dual and extremum. Following Noble & Sewell (1972, 1976) one considers the equations obtained by setting the partial derivatives of X with respect to $\hat{D}_{\mathrm{I}}$ and $\hat{D}_{\mathrm{II}}$ equal to zero. For the linear equation (3.3), this leads to the system of equations

$$X,_{\mathrm{I}}(\hat{u}) = \hat{P}_{\mathrm{I}}\hat{u} - \hat{f}_{\mathrm{I}} = 0 \tag{3.7a}$$
$$X,_{\mathrm{II}}(\hat{u}) = \hat{P}_{\mathrm{II}}\hat{u} - \hat{f}_{\mathrm{II}} = 0. \tag{3.7b}$$

These equations, which are obviously equivalent to (3.3), can also be derived from the latter equation by taking projections on $\hat{D}_{\mathrm{I}}$ and $\hat{D}_{\mathrm{II}}$.

THEOREM 3.4. *Let $\hat{P}\colon \hat{D} \to \hat{D}^*$ be saddle with respect to the decomposition $\hat{D}_{\mathrm{I}}$, $\hat{D}_{\mathrm{II}}$ of $\hat{D}$. Define the affine subspaces*

$$\hat{\mathscr{D}}_a = \{\hat{u} \in \hat{D} \mid \hat{u} \text{ satisfies } (3.7\text{a})\} \tag{3.8a}$$

$$\hat{\mathscr{D}}_b = \{\hat{u} \in \hat{D} \mid \hat{u} \text{ satisfies } (3.7\text{b})\}. \tag{3.8b}$$

Then:

(*i*) *$\hat{u} \in \hat{D}$ satisfies* (3.3) *if and only if $\hat{u} \in \hat{\mathscr{D}}_a \cap \hat{\mathscr{D}}_b$.*

(*ii*) *For every $\hat{u}_a \in \hat{\mathscr{D}}_a$ and $\hat{u}_b \in \hat{\mathscr{D}}_b$*

$$2[X(\hat{u}_b)-X(\hat{u}_a)] = \|\hat{u}_b-\hat{u}_a\|_{\hat{P}}^2 \geqslant 0. \tag{3.9}$$

(*iii*) *If the maximum of X on $\hat{\mathscr{D}}_a$ and the minimum on $\hat{\mathscr{D}}_b$ coincide and are attained at $\hat{U}_a$ and $\hat{U}_b$ respectively, then*

$$\hat{U}_a \approx_{\hat{P}} \hat{U}_b. \tag{3.10}$$

(*iv*) *If a solution $\hat{U} \in \hat{D}$ of* (3.3) *exists, then*

$$2[X(\hat{u}_b)-X(\hat{u}_a)] = \|\hat{u}_b-\hat{U}\|_{\hat{P}}^2+\|\hat{u}_a-\hat{U}\|_{\hat{P}}^2 \geqslant 0. \tag{3.11}$$

In this case the maximum of X on $\hat{\mathscr{D}}_a$ is attained at $\hat{u}_a$ if and only if $\hat{u}_a \approx_{\hat{P}} \hat{U}$, and the minimum of X on $\hat{\mathscr{D}}_b$ is attained at $\hat{u}_b$ if and only if $\hat{u}_b \approx_{\hat{P}} \hat{U}$. In particular if $\hat{U}$ and $\hat{V}$ are two solutions of (3.3), *they are $\hat{P}$-equivalent.*

Proof. Property (*i*) is obvious.

It is instructive to prove part (*iv*) before part (*ii*), because its proof is simpler. If $\hat{U} \in \hat{D}$ satisfies (3.3), and $\hat{u}_b \in \hat{\mathscr{D}}_b$, then

$$2X(\hat{U}) = \langle \hat{P}\hat{U}, \hat{U}\rangle - 2\langle \hat{f}, \hat{U}\rangle = -\langle \hat{P}\hat{U}, \hat{U}\rangle$$

$$2X(\hat{u}_b) = \langle \hat{P}\hat{u}_b, \hat{u}_b\rangle - 2\langle \hat{P}\hat{U}, \hat{u}_b\rangle.$$

Thus

$$2[X(\hat{u}_b)-X(\hat{U})] = \langle \hat{P}(\hat{u}_b-\hat{U}), \hat{u}_b-\hat{U}\rangle$$

because $\hat{P}$ is symmetric. Now $\hat{u}_b-\hat{U} \in \hat{D}_b$ (where $\hat{D}_b$ is defined by an extension of (2.9b)), in view of the fact that $\hat{U}$ as well as $\hat{u}_b$ satisfy (3.7b); thus, application of Lemma 2.5 (equation (2.10b)) leads to

$$2[X(\hat{u}_b)-X(\hat{U})] = \|\hat{u}_b-\hat{U}\|_{\hat{P}}^2 \geqslant 0. \tag{3.12a}$$

Similarly

$$2[X(\hat{U})-X(\hat{u}_a)] = \|\hat{u}_a-\hat{U}\|_{\hat{P}}^2 \geqslant 0. \tag{3.12b}$$

Adding these inequalities yields (3.11).

To prove property (*ii*), let $\hat{u}_a \in \hat{\mathscr{D}}_a$ and $\hat{u}_b \in \hat{\mathscr{D}}_b$: then

$$2X(\hat{u}_a) = \langle \hat{P}\hat{u}_a-2\hat{f}, \hat{u}_a\rangle = 2\langle \hat{P}\hat{u}_a-\hat{f}, \hat{u}_a\rangle - \langle \hat{P}\hat{u}_a, \hat{u}_a\rangle$$

$$= 2\langle \hat{P}\hat{u}_a-\hat{f}, \hat{u}_{a\mathrm{II}}\rangle - \langle \hat{P}\hat{u}_a, \hat{u}_a\rangle = \langle \hat{P}\hat{u}_{a\mathrm{II}}, \hat{u}_{a\mathrm{II}}\rangle - \langle \hat{P}\hat{u}_{a\mathrm{I}}, \hat{u}_{a\mathrm{I}}\rangle - 2\langle \hat{f}, \hat{u}_{a\mathrm{II}}\rangle$$

$$= -\|\hat{u}_a\|_{\hat{P}}^2 - 2\langle \hat{P}\hat{u}_b, \hat{u}_{a\mathrm{II}}\rangle = -\|\hat{u}_a\|_{\hat{P}}^2 - 2\langle \hat{P}\hat{u}_{b\mathrm{I}}, \hat{u}_{a\mathrm{II}}\rangle - 2\langle \hat{P}\hat{u}_{b\mathrm{II}}, \hat{u}_{a\mathrm{II}}\rangle \tag{3.13a}$$

where use has been made of the norm given before Definition 2.5, and of equations (3.7) and (3.8). Similarly

$$2X(\hat{u}_b) = \|\hat{u}_b\|_{\hat{P}}^2 - 2\langle \hat{P}\hat{u}_{a\mathrm{I}}, \hat{u}_{b\mathrm{I}}\rangle - 2\langle \hat{P}\hat{u}_{a\mathrm{II}}, \hat{u}_{b\mathrm{I}}\rangle. \tag{3.13b}$$

Relation (3.9) is obtained by subtracting (3.13a) from (3.13b).

Property (*iii*) is obviously implied by (*ii*).

4. General Theory of Non-Negative Unsymmetric Operators

Let $P: D \to D^*$ be a non-negative and not necessarily symmetric operator. Assume $f, g \in D^*$ are given functionals, and consider the system of equations

$$Pu_1 = f \tag{4.1a}$$

$$P^*u_2 = g \tag{4.1b}$$

where $u_1, u_2 \in D$. This system of equations may be written in terms of a symmetric operator $\hat{P}: \hat{D} \to \hat{D}^*$ where $\hat{D} = D \oplus D$ is the outer sum of D with itself, i.e. it is the space whose elements $\hat{u} = (u_1, u_2) \in \hat{D}$ are ordered pairs of elements $u_1, u_2 \in D$. To express equations (4.1) in the form

$$\hat{P}\hat{u} = \hat{f} \tag{4.2}$$

let $\hat{P}: \hat{D} \to \hat{D}^*$ and $\hat{f} \in \hat{D}^*$ be such that for every $\hat{u}, \hat{v} \in \hat{D}$

$$\langle \hat{P}\hat{u}, \hat{v} \rangle = \langle Pu_1, v_2 \rangle + \langle P^*u_2, v_1 \rangle \tag{4.3}$$

and

$$\langle \hat{f}, \hat{v} \rangle = \langle f, v_2 \rangle + \langle g, v_1 \rangle. \tag{4.4}$$

The operator $\hat{P}$ is clearly self-adjoint because

$$\langle \hat{P}\hat{u}, \hat{v} \rangle = \langle Pu_1, v_2 \rangle + \langle Pv_1, u_2 \rangle. \tag{4.5}$$

Due to this fact it is possible to formulate variational principles for the system (4.1). The procedure just described is a rigorous generalization of the mirror method (Morse & Feshbach, 1953).

THEOREM 4.1. *Let $X: \hat{D} \to R$ be given for every $\hat{u} \in \hat{D}$ by*

$$X(\hat{u}) = X(u_1, u_2) = \langle Pu_1, u_2 \rangle - \langle f, u_2 \rangle - \langle g, u_1 \rangle. \tag{4.6}$$

Then equations (4.1) *are satisfied if and only if*

$$X'(\hat{u}) = 0. \tag{4.7}$$

Proof. It follows from Theorem 3.3, because

$$X(\hat{u}) = \tfrac{1}{2}\langle \hat{P}\hat{u}, \hat{u} \rangle - \langle \hat{f}, \hat{u} \rangle. \tag{4.8}$$

The operator $\hat{P}$ does not have the property of being non-negative or non-positive because

$$\langle \hat{P}\hat{u}, \hat{u} \rangle = 2\langle Pu_1, u_2 \rangle. \tag{4.9}$$

In general when a self-adjoint operator which is functional-valued and linear is given, it is possible to find decompositions $\hat{D}_{\mathrm{I}}$ and $\hat{D}_{\mathrm{II}}$ of the space, such that the operator is saddle with respect to them (Herrera & Bielak, 1976). Finding such decompositions may be difficult in some instances; however, for the operator $\hat{P}$ here considered this is achieved easily.

LEMMA 4.2. *Let α, β, γ, δ be real numbers such that $\alpha\delta - \beta\gamma \neq 0$, $\alpha\beta < 0$, and $\gamma\delta > 0$. Define*

$$\hat{D}_{\mathrm{I}} = \{\hat{u} \in \hat{D} \mid \alpha u_1 + \beta u_2 = 0\} \tag{4.10a}$$

$$\hat{D}_{\mathrm{II}} = \{\hat{u} \in \hat{D} \mid \gamma u_1 + \delta u_2 = 0\}. \tag{4.10b}$$

Then:

(*i*) *The subspaces $\hat{D}_{\mathrm{I}}$, $\hat{D}_{\mathrm{II}}$ constitute a decomposition of $\hat{D}$. The projections $\hat{u}_{\mathrm{I}}$ and $\hat{u}_{\mathrm{II}}$ of any $\hat{u} \in \hat{D}$ are*

$$\hat{u}_{\mathrm{I}} = \lambda(-\beta\gamma u_1 - \beta\delta u_2, \alpha\gamma u_1 + \alpha\delta u_2) \tag{4.11a}$$

$$\hat{u}_{\mathrm{II}} = \lambda(\alpha\delta u_1 + \beta\delta u_2, -\alpha\gamma u_1 - \beta\gamma u_2) \tag{4.11b}$$

$$\lambda = \frac{1}{\alpha\delta-\beta\gamma}. \tag{4.12}$$

(ii) *The operator $\hat{P}$ is saddle with respect to $\hat{D}_{\text{I}}$, $\hat{D}_{\text{II}}$. Even more, if P is strictly positive then $\hat{P}$ is strictly saddle.*

(iii) *Let*

$$\gamma_1 = -\beta/\alpha > 0, \qquad \gamma_2 = \gamma/\delta > 0 \tag{4.13a}$$

$$m_1 = \frac{1}{\frac{1}{2}(\gamma_1+1/\gamma_2)} > 0, \qquad m_2 = \frac{1}{\frac{1}{2}(\gamma_2+1/\gamma_1)}. \tag{4.13b}$$

$$\|\hat{u}\|_{\hat{P}}^2 = m_1\|u_1\|_P^2+m_2\|u_2\|_P^2. \tag{4.14}$$

Thus, the relation $\hat{u} \approx_{\hat{P}} \hat{v}$ holds, if and only if the relations $u_1 \approx_P v_1$ and $u_2 \approx_P v_2$ are satisfied simultaneously.

(iv) *Each of the equations* (3.7a) *and* (3.7b) *is equivalent to*

$$\alpha Pu_1-\delta P^*u_2 = \alpha f-\beta g \tag{4.15a}$$

$$\gamma Pu_1-\delta P^*u_2 = \gamma f-\delta g \tag{4.15b}$$

respectively.

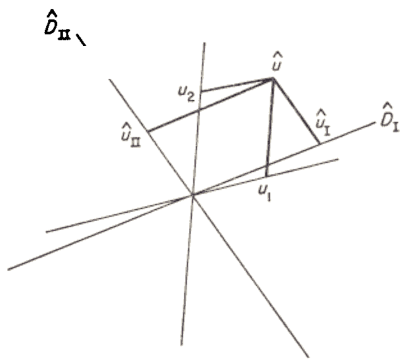

FIG. Decomposition of $\hat{D}$ to make $\hat{P}$ a saddle operator.

Proof. The decomposition of $\hat{D}$ is shown schematically in Fig. 1. The definitions (4.10) of $\hat{D}_{\text{I}}$ and $\hat{D}_{\text{II}}$ imply that these two sets are subspaces of $\hat{D}$. To prove that they constitute a decomposition of $\hat{D}$, it is only necessary to show that $\hat{u}_{\text{I}}$ and $\hat{u}_{\text{II}}$ as given by (4.11) have the following properties: (a) $\hat{u}_{\text{I}} \in \hat{D}_{\text{I}}$ and $\hat{u}_{\text{II}} \in \hat{D}_{\text{II}}$; (b) $\hat{u} = \hat{u}_{\text{I}}+\hat{u}_{\text{II}}$; (*c*) this is the only couple of elements of $\hat{D}$ that has properties (a) and (b). The proof of (a), (b) and (c) is straightforward.

In view of (4.9) and (4.10a), for every $\hat{u} \in \hat{D}_{\text{I}}$ one has

$$-\alpha\beta\langle \hat{P}\hat{u}, \hat{u}\rangle = \alpha^2\langle Pu_1, u_1\rangle+\beta^2\langle Pu_2, u_2\rangle \geqslant 0.$$

Similarly, for every $\hat{u} \in \hat{D}_{\text{II}}$

$$\gamma\delta\langle \hat{P}\hat{u}, \hat{u}\rangle = -\gamma^2\langle Pu_1, u_1\rangle-\delta^2\langle Pu_2, u_2\rangle \leqslant 0.$$

From this property (*ii*) follows easily.

Property (*iii*) is obtained by direct computations.

The definition of projections implies that equation (3.7a) is satisfied if and only if

$$\langle \hat{P}\hat{u}-\hat{f}, \hat{v}\rangle = 0 \tag{4.16}$$

for every $\hat{v} \in \hat{D}_{\text{I}}$. Using equations (4.3), (4.4) and (4.10a), it is seen that (4.16) is equivalent to

$$-\alpha\langle Pu_1-f, v_1\rangle+\beta\langle P^*u_2-g, v_1\rangle = 0 \tag{4.17}$$

for every $v_1 \in D$. Hence (4.15a). The proof of (4.15b) is similar.

In view of this lemma, Theorem 3.4 can be applied to this case.

THEOREM 4.3. *Let $X\colon \hat{D} \to R$ be given by* (4.6), $\alpha\delta-\beta\gamma \neq 0$, $\alpha\beta < 0$ *and* $\gamma\delta > 0$. *Define the affine subspaces:*

$$\hat{\mathscr{D}}_a = \{\hat{u} \in \hat{D} \mid \hat{u} \text{ satisfies } (4.15\text{a})\} \tag{4.18a}$$

$$\hat{\mathscr{D}}_b = \{\hat{u} \in \hat{D} \mid \hat{u} \text{ satisfies } (4.15\text{b})\}. \tag{4.18b}$$

Then:

(*i*) $\hat{u} = (u_1, u_2) \in \hat{D}$ *is a solution of equations* (4.1) *if and only if* $\hat{u} \in \hat{\mathscr{D}}_a \bigcap \hat{\mathscr{D}}_b$.

(*ii*) *For every* $\hat{u}_a = (u_{a1}, u_{a2}) \in \hat{\mathscr{D}}_a$ *and every* $\hat{u}_b = (u_{b1}, u_{b2}) \in \hat{\mathscr{D}}_b$

$$2[X(\hat{u}_b)-X(\hat{u}_a)] = m_1\|u_{a1}-u_{b1}\|_P^2+m_2\|u_{a2}-u_{b2}\|_P^2 \geqslant 0. \tag{4.19}$$

(*iii*) *If the maximum of X on $\hat{\mathscr{D}}_a$ and the minimum on $\hat{\mathscr{D}}_b$ coincide and are attained at $\hat{u}_a$ and $\hat{u}_b$ respectively*

$$u_{a1} \approx_P u_{b1} \tag{4.20a}$$

and simultaneously

$$u_{a2} \approx_P u_{b2}. \tag{4.20b}$$

(*iv*) *If a solution $\hat{U} = (U_1, U_2) \in \hat{D}$ of* (4.1) *exists, then*

$$2[X(\hat{u}_b)-X(\hat{u}_a)] = m_1\{\|u_{b1}-U_1\|_P^2+\|u_{a1}-U_1\|_P^2\}+ m_2\{\|u_{b2}-U_2\|_P^2+\|u_{a2}-U_2\|_P^2\} \tag{4.21}$$

for every $\hat{u}_a \in \hat{\mathscr{D}}_a$ and $\hat{u}_b \in \hat{\mathscr{D}}_b$.

Proof. Using Lemma 4.2, the proof of this theorem is a straightforward application of Theorem 3.4.

5. Extension to Problems Defined in an Affine Subspace

In many applications to partial differential equations, the operator involved may be saddle or non-negative when attention is restricted to functions satisfying certain boundary conditions. In such problems the admissible functions usually constitute an affine subspace and the theory would be unduly restricted if such cases were not included.

A set $\hat{\mathscr{E}} \subset \hat{D}$ is said to be an affine subspace if there is a subspace $\hat{E} \subset \hat{D}$ and an element $\hat{w} \in \hat{D}$ such that $\hat{\mathscr{E}} = \hat{w}+\hat{E}$. Clearly, $\hat{E} = \hat{\mathscr{E}}$ when $\hat{w} \in \hat{E}$, so that a linear subspace is always an affine subspace.

Given a functional $X\colon \hat{D} \to R$, for each $\hat{u} \in \hat{\mathscr{E}}$ the variation of X at $\hat{u}$ is a linear functional $\delta X(\hat{u}) \in \hat{E}^*$, such that

$$\langle \delta X(\hat{u}), \hat{v}\rangle = \langle X'(\hat{u}), \hat{v}\rangle \tag{5.1}$$

for every $\hat{v} \in \hat{E}$. When a decomposition $\hat{E}_{\text{I}}, \hat{E}_{\text{II}}$ of $\hat{E}$ is available, at every $\hat{u} \in \hat{\mathscr{E}}$ the partial variations $\delta_i X(\hat{u}) \in \hat{E}^*$ ($i = \text{I, II}$) are defined in a manner similar to partial derivatives (Herrera & Bielak, 1974).

Variational principles which hold when the set of admissible functions is $\hat{D}$ itself, but which remain valid when the set of admissible functions is restricted to be an affine subspace $\hat{\mathscr{E}} \subset \hat{D}$, are known in many instances. The validity of such a procedure frequently depends on the following property of the affine subspace (Herrera & Bielak, 1976).

Definition 5.1. Let $\hat{P}: \hat{D} \to \hat{D}^*$ and $\hat{f} \in \hat{D}^*$ be given. An affine subspace $\hat{\mathscr{E}} = \hat{w}+\hat{E}$ is said to be determinative for problem (4.2) if it possesses the following property: whenever $\hat{u} \in \hat{\mathscr{E}}$, the fact that

$$\langle \hat{P}\hat{u}-\hat{f}, \hat{v}\rangle = 0 \quad \text{for every } \hat{v} \in \hat{E} \tag{5.2}$$

implies (4.2).

LEMMA 5.2. *Let $\hat{P}: \hat{D} \to \hat{D}^*$ be self-adjoint, $X(\hat{u})$ be given by* (3.4) *and $\hat{\mathscr{E}} \subset \hat{D}$ be an affine subspace determinative for problem* (3.3). *Then $\hat{u} \in \hat{\mathscr{E}}$ satisfies* (3.3) *if and only if*

$$\delta X(\hat{u}) = 0. \tag{5.3}$$

If in addition $\hat{E}_{\mathrm{I}}, \hat{E}_{\mathrm{II}}$ is a decomposition of $\hat{E}$, then (5.3) *can be replaced by:*

$$\delta_{\mathrm{I}} X(\hat{u}) = 0 \tag{5.4a}$$

$$\delta_{\mathrm{II}} X(\hat{u}) = 0. \tag{5.4b}$$

Proof. Equation (5.3) means

$$\langle \hat{P}\hat{u}-\hat{f}, \hat{v}\rangle = \langle X'(\hat{u}), \hat{v}\rangle = 0 \tag{5.5}$$

whenever $\hat{v} \in \hat{E}$. By the definition of determinative affine subspace the first part of the lemma follows. The other part is obvious.

The following result can now be easily derived from Theorem 3.4.

THEOREM 5.3. *Assume $\hat{\mathscr{E}} = \hat{w}+\hat{E}$ is an affine subspace, determinative for problem* (3.3). *Assume in addition that $\hat{P}: \hat{D} \to \hat{D}^*$ is saddle on $\hat{E}$ with respect to the decomposition $\hat{E}_{\mathrm{I}}, \hat{E}_{\mathrm{II}}$. The properties* (i) *to* (iv) *of Theorem* 3.4 *remain valid when the sets $\hat{\mathscr{D}}_a$ and $\hat{\mathscr{D}}_b$ (equations* (3.8)) *are replaced respectively by:*

$$\hat{\mathscr{E}}_a = \{\hat{u} \in \hat{\mathscr{E}} \mid \hat{u} \text{ satisfies } (5.4a)\} \tag{5.6a}$$

$$\hat{\mathscr{E}}_b = \{\hat{u} \in \hat{\mathscr{E}} \mid \hat{u} \text{ satisfies } (5.4b)\}. \tag{5.6b}$$

Proof. Under the assumptions of the theorem and writing $\hat{u} = \hat{w}+\hat{v}$, problem (3.3) is equivalent to

$$\tilde{P}\hat{v} = \tilde{f}-\hat{P}\hat{w} \tag{5.7}$$

where $\hat{v} \in \hat{E}$, while $\tilde{P}: \hat{E} \to \hat{E}^*$ and $\tilde{f} \in \hat{E}^*$ are the restrictions of $\hat{P}$ and $\hat{f}$ to $\hat{E}$. The theorem follows now by application of Theorem 3.4 to equation (5.7), after observing that

$$\tfrac{1}{2}\langle \tilde{P}\hat{v}, \hat{v}\rangle - \langle \tilde{f}-\hat{P}\hat{w}, \hat{v}\rangle = X(\hat{u}) + \langle \hat{f}, \hat{w}\rangle - \tfrac{1}{2}\langle \hat{P}\hat{w}, \hat{w}\rangle. \tag{5.8}$$

That is, the functional $X(\hat{u})$ differs from the functional associated with the problem (5.7) formulated on $\hat{E}$ by a constant term only.

THEOREM 5.4. *Assume $P: D \to D^*$ is non-negative on the subspace $E \subset D$. Let $\mathscr{E}_1 = w_1+E$ and $\mathscr{E}_2 = w_2+E$, with $w_1, w_2 \in D$, be two affine subspaces of D such that they are determinative for problems* (4.1a) *and* (4.1b) *respectively. Define the affine subspace $\hat{\mathscr{E}} \subset \hat{D}$ by $\hat{\mathscr{E}} = \mathscr{E}_1 \oplus \mathscr{E}_2 = \hat{w}+\hat{E}$, where $\hat{w} = w_1 \oplus w_2 = (w_1, w_2)$, and $\hat{E} = E \oplus E$. Let*

$$\hat{\mathscr{E}}_a = \{\hat{u} \in \hat{\mathscr{E}} \mid \hat{u} \text{ satisfies } (5.4a)\} \tag{5.9a}$$

$$\hat{\mathscr{E}}_b = \{\hat{u} \in \hat{\mathscr{E}} \mid \hat{u} \text{ satisfies } (5.4b)\}. \tag{5.9b}$$

Then Theorem 4.3 *remains true if the sets* $\hat{\mathscr{D}}_a$ *and* $\hat{\mathscr{D}}_b$ *are replaced by sets* $\hat{\mathscr{E}}_a$ *and* $\hat{\mathscr{E}}_b$ *respectively.*

Proof. This follows from Theorem 5.3, observing that the affine subspace $\hat{\mathscr{E}}$ is determinative for problem 4.2, while $\hat{P}$ is saddle on $\hat{E} = E \oplus E$ with respect to the decomposition $\hat{E}_{\mathrm{I}}$, $\hat{E}_{\mathrm{II}}$ of $\hat{E}$. Here

$$\hat{E}_{\mathrm{I}} = \{\hat{v} = (v_1, v_2) \in \hat{E} \mid \alpha v_1 + \beta v_2 = 0\} \tag{5.10a}$$

and

$$\hat{E}_{\mathrm{II}} = \{\hat{v} = (v_1, v_2) \in \hat{E} \mid \gamma v_1 + \delta v_2 = 0\}. \tag{5.10b}$$

6. Applications

In the applications to be considered, G will be a region in n-dimensional euclidean space, with closure $\bar{G}$ and boundary $S = S_1 \bigcup S_2$, where S_1 and S_2 are disjoint. The linear space D of functions will be assumed to be such that the differential operators and integrals are well defined (possibly in a generalized sense). Such would be the case if the functions are C^2 on $\bar{G} \times [0, T]$; this is however, unnecessarily restrictive.

A. The Heat Equation

For every $u, v \in D$, the operator $P\colon D \to D^*$ will be defined by:

$$\langle Pu, v\rangle = \int_0^T \left\{\int_G \frac{\partial v}{\partial t}\left(\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x_i\, \partial x_i}\right) dx + \int_{S_1} \frac{\partial u}{\partial t}\frac{\partial v}{\partial n}\, dx + \int_{S_2} \frac{\partial u}{\partial n}\frac{\partial v}{\partial t}\, dx\right\} dt + \int_G \left[\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}\right]_{t=0} dx. \tag{6.1}$$

Using integration by parts, a convenient expression for the adjoint operator is obtained:

$$\langle P^* u, v\rangle = \int_0^T \left\{\int_G \frac{\partial v}{\partial t}\left(\frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial x_i\, \partial x_i}\right) dx - \int_{S_1} \frac{\partial u}{\partial t}\frac{\partial v}{\partial n}\, dx - \int_{S_2} \frac{\partial u}{\partial n}\frac{\partial v}{\partial t} dx\right\} dt + \int_G \left[\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}\right]_{t=T} dx. \tag{6.2}$$

Therefore:

$$\langle Pu, u\rangle = \int_0^T \int_G \left(\frac{\partial u}{\partial t}\right)^2 dx\, dt + \frac{1}{2}\int_G \left\{\left[\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_i}\right]_{t=0} + \left[\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_i}\right]_{t=T}\right\} dx = \|u\|_P^2. \tag{6.3}$$

Assume f_G, f_{S1} and f_{S2} are continuous functions defined on $\bar{G} \times [0, T]$, $S_1 \times [0, T]$ and $S_2 \times [0, T]$ respectively, and let f_0 be C^1 on $\bar{G}$. Define $f \in D^*$, for every $v \in D$, by:

$$\langle f, v\rangle = \int_0^T \left\{\int_G f_G \frac{\partial v}{\partial t}\, dx + \int_{S_1} f_{S1} \frac{\partial v}{\partial n}\, dx + \int_{S_2} f_{S2} \frac{\partial v}{\partial t}\, dx\right\} dt + \int_G \left[\frac{\partial f_0}{\partial x_i}\frac{\partial v}{\partial x_i}\right]_{t=0} dx. \tag{6.4}$$

Let g_G, g_{S1}, g_{S2} and g_T be another set of functions satisfying the same hypotheses and define $g \in D^*$ by:

$$\langle g, v\rangle = \int_0^T \left\{\int_G g_G \frac{\partial v}{\partial t}\, dx - \int_{S_1} g_{S1} \frac{\partial v}{\partial n}\, dx - \int_{S_2} g_{S2} \frac{\partial v}{\partial t}\, dx\right\} dt + \int_G \left[\frac{\partial g_T}{\partial x_i}\frac{\partial v}{\partial x_i}\right]_{t=T} dx.$$

Application of the so-called "fundamental lemma of the calculus of variations" shows that for functions $u_1 \in D$, the equation

$$Pu_1 = f \tag{6.6}$$

is equivalent to the system:

$$\frac{\partial u_1}{\partial t} - \frac{\partial^2 u_1}{\partial x_i \partial x_i} = f_G \quad \text{on } \bar{G} \times [0, T]$$

$$\frac{\partial u_1}{\partial t} = f_{S1} \quad \text{on } S_1 \times [0, T]$$

$$\frac{\partial u_1}{\partial n} = f_{S2} \quad \text{on } S_2 \times [0, T] \tag{6.7c}$$

$$\frac{\partial u_1}{\partial x_i} = \frac{\partial f_0}{\partial x_i} \quad \text{on } \bar{G} \text{ at } t = 0. \tag{6.7d}$$

Similarly, for functions $u_2 \in D$, the equation

$$P^* u_2 = g$$

is equivalent to the problem

$$\frac{\partial u_2}{\partial t} + \frac{\partial^2 u_2}{\partial x_i \partial x_i} = g_G \quad \text{on } \bar{G} \times [0, T] \tag{6.9a}$$

$$\frac{\partial u_2}{\partial t} = g_{S1} \quad \text{on } S_1 \times [0, T] \tag{6.9b}$$

$$\frac{\partial u_2}{\partial n} = g_{S2} \quad \text{on } S_2 \times [0, T]$$

$$\frac{\partial u_2}{\partial x_i} = \frac{\partial g_T}{\partial x_i} \quad \text{on } \bar{G} \text{ at } t = T. \tag{6.9d}$$

A more standard form of these problems for the heat equation is obtained by replacing equations (6.7d) and (6.9d) by

$$u_1 = f_0 \quad \text{on } \bar{G} \text{ at } t = 0 \tag{6.10}$$

and

$$u_2 = g_T \quad \text{on } \bar{G} \text{ at } t = T \tag{6.11}$$

respectively. When these equations hold, the corresponding problems will be called initial and terminal value problems respectively. Observe that a solution of problem (6.7) or (6.9) may differ from those of the latter by at most a constant.

The space $\hat{D}$ and the operator $\hat{P}$: $\hat{D} \to \hat{D}^*$ is defined as in Section 4. The functional X given by equation (4.6) is:

$$\begin{aligned} X(u_1, u_2) = \int_0^T \Bigg\{ &\int_G \left[\frac{\partial u_2}{\partial t} \left(\frac{\partial u_1}{\partial t} - \frac{\partial^2 u_1}{\partial x_i \partial x_i} - f_G \right) - g_G \frac{\partial u_1}{\partial t} \right] d\mathbf{x} + \\ &\int_{S1} \left[\left(\frac{\partial u_1}{\partial t} - f_{S1} \right) \frac{\partial u_2}{\partial n} - g_{S1} \frac{\partial u_1}{\partial n} \right] d\mathbf{x} + \\ &\int_{S2} \left[\left(\frac{\partial u_1}{\partial n} - f_{S2} \right) \frac{\partial u_2}{\partial t} - g_{S2} \frac{\partial u_1}{\partial t} \, d\mathbf{x} \right] dt + \\ &\int_G \left\{ \left[\left(\frac{\partial u_1}{\partial x_i} - \frac{\partial f_0}{\partial x_i} \right) \frac{\partial u_2}{\partial x_i} \right]_{t=0} - \left[\frac{\partial g_T}{\partial x_i} \frac{\partial u_1}{\partial x_i} \right]_{t=T} \right\} dx. \end{aligned} \tag{6.12}$$

Thus:

THEOREM 6.1 *The functions $u_1, u_2 \in D$ are solutions of problems* (6.7) *and* (6.9) *if and only if*

$$X'(u_1, u_2) = 0.$$

There are several alternative ways of stating the dual extremum principles derived from Theorem 4.3. One that we have found convenient is the following:

THEOREM 6.2. *For every $v \in D$, define*:

$$u_{a1}(\mathbf{x}, t) = f_0(\mathbf{x})+\frac{1}{2\alpha}\left\{v(\mathbf{x}, t)-v(\mathbf{x}, 0)+\int_0^t\left(\frac{\partial^2 v}{\partial x_i\,\partial x_i}+\alpha f_G-\beta g_G\right)dt\right\} \tag{6.13a}$$

$$u_{a2}(\mathbf{x}, t) = \frac{1}{\beta}(v-\alpha u_{1a}). \tag{6.13b}$$

Let the set of functions $A \subset D$ be such that $v \in A$ implies $u_{a1} \in D$ and in addition:

$$\frac{\partial^2}{\partial x_i\,\partial x_i}\int_0^T v(\mathbf{x}, t)\,dt = v(T)+v(0)-2\alpha f_0+2\beta g_T+\int_0^T(\beta g_G-\alpha f_G)\,dt; \quad \mathbf{x} \in \bar{G} \tag{6.14a}$$

$$\frac{\partial v}{\partial t} = \alpha f_{S1}+\beta g_{S1} \quad on\ S_1\times[0, T] \tag{6.14b}$$

$$\frac{\partial v}{\partial n} = \alpha f_{S2}+\beta g_{S2} \quad on\ S_2\times[0, T]. \tag{6.14c}$$

Assume the functional Ω_a: $A \to R$ is given by:

$$\Omega_a(v) = \frac{\alpha}{\beta}\left\{\int_0^T\int_G\left(\frac{\partial u_{1a}}{\partial t}\right)^2 d\mathbf{x}\,dt+\frac{1}{2}\int_G\left[\frac{\partial u_{1a}}{\partial x_i}\frac{\partial u_{1a}}{\partial x_i}\right]_{t=T} d\mathbf{x}+\frac{1}{2}\int_G\frac{\partial f_0}{\partial x_i}\frac{\partial f_0}{\partial x_i}d\mathbf{x}\right\}-$$
$$\frac{1}{\beta}\int_0^T\left\{\int_G f_G\frac{\partial v}{\partial t}d\mathbf{x}+\int_{S1} f_{S1}\frac{\partial v}{\partial n}d\mathbf{x}+\int_{S2} f_{S2}\frac{\partial v}{\partial t}d\mathbf{x}\right\}dt+$$
$$\int_G\left[\frac{\partial f_0}{\partial x_i}\frac{\partial v}{\partial x_i}\right]_{t=0} d\mathbf{x}. \tag{6.15}$$

The definitions of u_{b1}, u_{b2}, $B \subset D$ and Ω_b: $B \to R$, are obtained replacing α by γ and β by δ in the above definitions.

Let $v \in A$ and $w \in B$. Write u_{a1}, u_{a2} for the functions associated with v by equations (6.13) *and u_{b1}, u_{b2} for the functions associated with w by the modified version of those equations. Then*:

(*i*) *$u_{a1} \equiv u_{b1}$ and simultaneously $u_{a2} \equiv u_{b2}$, if and only if u_{a1}, u_{a2} are solutions respectively of the initial and terminal value problems for the heat equation.*

(*ii*)

$$2[\Omega_b(w)-\Omega_a(v)] = m_1\|u_{a1}-u_{b1}\|_P^2+m_2\|u_{a2}-u_{b2}\|_P^2 \geqslant 0 \tag{6.16}$$

where $\| \|_P^2$ is given by (6.3).

Assume the maximum of Ω_a and the minimum of Ω_b coincide and are attained at v and w respectively. Then

$$u_{a1} \equiv u_{b1} \quad and \quad u_{a2} \equiv u_{b2}. \tag{6.17}$$

Thus, they are solutions of the initial and terminal value problems for the heat equation.

If solutions U_1, $U_2 \in D$ of the initial and terminal value problems respectively exist, then

$$2[\Omega_b(w)-\Omega_a(v)] = m_1\{\|u_{b1}-U_1\|_P^2+\|u_{a1}-U_1\|_P^2\}+ m_2\{\|u_{b2}-U_2\|_P^2+\|u_{a2}-U_2\|_P^2\} \geqslant 0. \tag{6.18}$$

Proof. To apply Theorem 4.3, observe that in view of equations (6.1), (6.2), (6.4) and (6.5), equation (4.15a) is equivalent to:

$$\alpha\frac{\partial u_1}{\partial t}-\beta\frac{\partial u_2}{\partial t}-\alpha\frac{\partial^2 u_1}{\partial x_i\,\partial x_i}-\beta\frac{\partial^2 u_2}{\partial x_i\,\partial x_i} = \alpha f_G-\beta g_G \quad \text{on } \bar{G}\times[0, T] \tag{6.19a}$$

$$\alpha\frac{\partial u_1}{\partial t}+\beta\frac{\partial u_2}{\partial t} = \alpha f_{S1}+\beta g_{S1} \quad \text{on } S_1\times[0, T] \tag{6.19b}$$

$$\alpha\frac{\partial u_1}{\partial n}+\beta\frac{\partial u_2}{\partial n} = \alpha f_{S2}+\beta g_{S2} \quad \text{on } S_2\times[0, T] \tag{6.19c}$$

$$\frac{\partial u_1}{\partial x_i} = \frac{\partial f_0}{\partial x_i} \quad \text{on } \bar{G} \text{ at } t = 0 \tag{6.19d}$$

$$\frac{\partial u_2}{\partial x_i} = \frac{\partial g_T}{\partial x_i} \quad \text{on } \bar{G} \text{ at } t = T. \tag{6.19e}$$

Let $v \in A$; then a straightforward computation shows that $u_{a1}, u_{a2} \in D$ as given by (6.13) satisfy equations (6.19a–d). On the other hand, equation (6.19e) is also satisfied by u_{a2}, by virtue of (6.14a).

In a similar fashion, it is shown that given $w \in B$, u_{b1} and u_{b2} satisfy equations (4.15b). Assertion (i) follows from Theorem 4.3 and the fact that the u functions satisfy equations (6.10) and (6.11).

When equation (4.15a) is satisfied the functional X, given by (4.6), becomes

$$X(u_1, u_2) = \frac{\alpha}{\beta}\langle Pu_1, u_1\rangle-\frac{1}{\beta}\langle f, \alpha u_1+\beta u_2\rangle. \tag{6.20a}$$

Similarly, when equation (4.15b) is satisfied

$$X(u_1, u_2) = \frac{\gamma}{\delta}\langle Pu_1, u_1\rangle-\frac{1}{\delta}\langle f, \delta u_1+\gamma u_2\rangle. \tag{6.20b}$$

In view of these facts and property (*i*), Theorem 4.3 yields the rest of the theorem.

Dual variational principles for diffusion equations were first obtained by Herrera (1974; Herrera & Bielak, 1976). Later Collins (1976) extended Herrera's results to a more general class of dissipative systems. Independently Brezis & Ekeland (1976) have obtained such principles for a class of parabolic equations. For the heat equation, Theorem 6.2 represents a definite improvement over previous results because these require solving Laplace's equation at every time t. Brezis & Ekeland's (1976) principles are in addition restricted to the case where S_2 is void and u vanishes identically on S.

Dual variational principles for the heat equation that have been derived in the past (Herrera, 1974; Herrera & Bielak, 1976; Collins, 1976; Brezis & Ekeland, 1976) can also be obtained using the general theory presented here. For this purpose, the operator $P: D \to D^*$ may be defined by:

$$\langle Pu, v\rangle = \int_0^T \left\{ \int_G v\left(\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x_i\, \partial x_i}\right) d\mathbf{x} + \int_{S1} u \frac{\partial v}{\partial n}\, d\mathbf{x} - \int_{S2} v \frac{\partial u}{\partial n}\, d\mathbf{x} \right\} dt + \int_G vu\, |_{t=0}\, d\mathbf{x} \tag{6.21}$$

which is positive on the linear subspace of functions with vanishing normal derivatives on S_2. The extension of the theory to problems formulated in affine subspaces, developed in Section 5, can then be applied to this operator to obtain those results. However, the details will not be carried out here.

B. The Wave Equation

For every $u, v \in D$, the operator $P: D \to D^*$ will be defined by:

$$\langle Pu, v\rangle = \int_0^T \left\{ \int_G \frac{\partial v}{\partial t}\left(\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x_i\, \partial x_i}\right) d\mathbf{x} + \int_{S1} \frac{\partial u}{\partial t} \frac{\partial v}{\partial n}\, d\mathbf{x} + \int_{S2} \frac{\partial u}{\partial n} \frac{\partial v}{\partial t}\, d\mathbf{x} \right\} dt + \int_G \left[\frac{\partial u}{\partial t}\frac{\partial v}{\partial t} + \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}\right]_{t=0} d\mathbf{x}.$$

Integration by parts shows that

$$\langle P^*u, v\rangle = -\int_0^T \left\{ \int_G \frac{\partial v}{\partial t}\left(\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x_i\, \partial x_i}\right) d\mathbf{x} + \int_{S1} \frac{\partial u}{\partial t} \frac{\partial v}{\partial n}\, d\mathbf{x} + \int_{S2} \frac{\partial u}{\partial n} \frac{\partial v}{\partial t}\, d\mathbf{x} \right\} dt + \int_G \left[\frac{\partial u}{\partial t}\frac{\partial v}{\partial t} + \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}\right] d\mathbf{x}.$$

The operator P is non-negative because

$$\langle Pu, u\rangle = \|u\|_P^2 = \frac{1}{2}\int_G \left\{ \left[\left(\frac{\partial u}{\partial t}\right)^2 + \frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_i}\right]_{t=0} + \left[\left(\frac{\partial u}{\partial t}\right)^2 + \frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_i}\right]_{t=T} \right\} d\mathbf{x}. \tag{6.24}$$

In addition to the functions f and g introduced in subsection A, let f_t and g_t be continuous functions defined on $\bar{G}$. Define $f \in D^*$ for every $v \in D$ by:

$$\langle f, v\rangle = \int_0^T \left\{ \int_G f_G \frac{\partial v}{\partial t}\, d\mathbf{x} + \int_{S1} f_{S1} \frac{\partial v}{\partial n}\, d\mathbf{x} + \int_{S2} f_{S2} \frac{\partial v}{\partial t}\, d\mathbf{x} \right\} dt + \int_G \left[f_t \frac{\partial v}{\partial t} + \frac{\partial f_0}{\partial x_i}\frac{\partial v}{\partial x_i}\right]_{t=0} d\mathbf{x} \tag{6.25}$$

and $g \in D^*$ by

$$\langle g, v\rangle = -\int_0^T \left\{ \int_G g_G \frac{\partial v}{\partial t}\, d\mathbf{x} + \int_{S1} g_{S1} \frac{\partial v}{\partial n}\, d\mathbf{x} + \int_{S2} g_{S2} \frac{\partial v}{\partial t}\, d\mathbf{x} \right\} dt + \int_G \left[g_t \frac{\partial v}{\partial t} + \frac{\partial g_T}{\partial x_i}\frac{\partial v}{\partial x_i}\right]_{t=T} d\mathbf{x}.$$

Then, applying the fundamental lemma of the calculus of variations, it is seen that the equation

$$Pu_1 = f$$

for $u_1 \in D$ is equivalent to the problem:

$$\frac{\partial^2 u_1}{\partial t^2} - \frac{\partial^2 u_1}{\partial x_i \partial x_i} = f_G \quad \text{on } \bar{G}\times[0, T] \tag{6.28a}$$

$$\frac{\partial u_1}{\partial t} = f_{S1} \quad \text{on } S_1\times[0, T]$$

$$\frac{\partial u_1}{\partial n} = f_{S2} \quad \text{on } S_2\times[0, T] \tag{6.28c}$$

$$\frac{\partial u_1}{\partial t} = f_t \quad \text{on } \bar{G} \text{ and } t = 0 \tag{6.28d}$$

$$\frac{\partial u_1}{\partial x_i} = \frac{\partial f_0}{\partial x_i} \quad \text{on } \bar{G} \text{ and } t = 0. \tag{6.28e}$$

Similarly, the equation

$$P^* u_2 = g$$

is equivalent to the problem:

$$\frac{\partial^2 u_2}{\partial t^2} - \frac{\partial^2 u_2}{\partial x_i \partial x_i} \quad \text{on } \bar{G}\times[0, T] \tag{6.30a}$$

$$\frac{\partial u_2}{\partial t} = g_{S1} \quad \text{on } S_1\times[0, T] \tag{6.30b}$$

$$\frac{\partial u_2}{\partial n} = g_{S2} \quad \text{on } S_2\times[0, T] \tag{6.30c}$$

$$\frac{\partial u_2}{\partial t} = g_t \quad \text{on } \bar{G} \text{ and } t = T \tag{6.30d}$$

$$\frac{\partial u_2}{\partial x_i} = \frac{\partial g_T}{\partial x_i} \quad \text{on } \bar{G} \text{ and } t = T. \tag{6.30e}$$

As in the case of the heat equation, the initial and terminal value problems are defined replacing (6.28e) and (6.30e) by equations (6.10) and (6.11) respectively; the observation made there about the relation between their solutions also applies here.

The space $\hat{D}$ and the operator $\hat{P}$: $\hat{D} \to \hat{D}^*$ will now be defined as in Section 4. The functional X given by equation (4.6) is

$$\begin{aligned} X(u_1, u_2) = \int_0^T \Bigg\{ & \int_G \left[\frac{\partial u_2}{\partial t}\left(\frac{\partial^2 u_1}{\partial t^2} - \frac{\partial^2 u_1}{\partial x_i \partial x_i} - f_G\right) + g_G \frac{\partial u_1}{\partial t}\right] d\mathbf{x} + \\ & \int_{S1} \left[\left(\frac{\partial u_1}{\partial t} - f_{S1}\right)\frac{\partial u_2}{\partial n} + g_{S1}\frac{\partial u_1}{\partial n}\right] d\mathbf{x} + \\ & \int_{S2} \left[\left(\frac{\partial u_1}{\partial n} - f_{S2}\right)\frac{\partial u_2}{\partial t} + g_{S2}\frac{\partial u_1}{\partial t}\right] d\mathbf{x} \Bigg\} dt + \\ & \int_G \left[\left(\frac{\partial u_1}{\partial t} - f_t\right)\frac{\partial u_2}{\partial t} + \left(\frac{\partial u_1}{\partial x_i} - \frac{\partial f_0}{\partial x_i}\right)\frac{\partial u_2}{\partial x_i}\right]_{t=0} d\mathbf{x} - \\ & \int_G \left(g_t \frac{\partial u_1}{\partial t} + \frac{\partial g_T}{\partial x_i}\frac{\partial u_1}{\partial x_i}\right)_{t=T} d\mathbf{x}. \end{aligned}$$

Theorem 4.1 yields a variational principle for the problem of the wave equation considered here.

THEOREM 6.3. *The functions* $u_1, u_2 \in D$ *are solutions of problems* (6.27) *and* (6.29) *simultaneously if and only if* $X'(\mathbf{u}) \equiv X'(u_1, u_2) = 0$.

Assume $\alpha, \beta, \gamma, \delta$ are real numbers satisfying the hypotheses of Lemma 4.2, and define $\hat{D}_{\mathrm{I}}$ and $\hat{D}_{\mathrm{II}}$ by (4.10). Let

$$\alpha' = \delta/\Delta; \qquad \beta' = -\beta/\Delta \tag{6.32a}$$

and

$$\gamma' = -\gamma/\Delta; \qquad \delta' = \alpha/\Delta \tag{6.32b}$$

where

$$\Delta = \alpha\delta - \beta\gamma. \tag{6.32c}$$

Given any $w_1, w_2 \in D$, define

$$u_1 = \alpha' w_1 + \beta' w_2 \tag{6.33a}$$

$$u_2 = \gamma' w_1 + \delta' w_2. \tag{6.33b}$$

Then these equations define a one to one mapping of $\hat{D}$ into itself whose inverse is

$$w_1 = \alpha u_1 + \beta u_2 \tag{6.34a}$$

$$w_2 = \gamma u_1 + \delta u_2. \tag{6.34b}$$

LEMMA 6.4. *Let* $w_1, w_2 \in D$, *and let* u_1, u_2 *be given by* (6.33). *Assume* w_1 *and* w_2 *are such that*

$$\frac{\partial u_1}{\partial t} = ft, \qquad u_1 = f_0 \quad \text{on } \bar{G} \text{ at } t = 0, \tag{6.35a}$$

$$\frac{\partial u_2}{\partial t} = gt, \qquad u_2 = g_T \quad \text{on } \bar{G} \text{ at } t = T. \tag{6.35b}$$

Then u_1 *and* u_2 *satisfy equation* (4.15a) *if and only if*:

$$\left.\begin{aligned} \frac{\partial^2 w_1}{\partial t^2} - \frac{\partial^2 w_1}{\partial x_i\, \partial x_i} &= \alpha f_G + \beta g_G \\ \frac{\partial w_1}{\partial t} &= \alpha f_{S1} + \beta g_{S1} \\ \frac{\partial w_1}{\partial n} &= \alpha f_{S2} + \beta g_{S2} \end{aligned}\right\} \tag{6.36a}$$

and equations (4.15b) *if and only if*:

$$\begin{aligned} \frac{\partial^2 w_2}{\partial t^2} - \frac{\partial^2 w_2}{\partial x_i\, \partial x_i} &= \gamma f_G + \delta g_G \\ \frac{\partial w_2}{\partial t} &= \gamma f_{S1} + \delta g_{S1} \\ \frac{\partial w_2}{\partial n} &= \gamma f_{S2} + \delta g_{S2}, \end{aligned} \tag{6.36b}$$

all on the corresponding domains.

Proof. By direct computation, using the definitions of P, f and g, it is seen that the assertion is true for functions satisfying equations (6.28d,e) and (6.30d,e). The lemma follows from the fact that this set of equations is satisfied whenever equations (6.35) hold.

THEOREM 6.5. *Let* $\Omega\colon \hat{D} \to R$ *be defined for every* $\hat{w} = (w_1, w_2) \in \hat{D}$ *by*

$$\Omega(\hat{w}) = \Omega(w_1, w_2) = X(u_1, u_2) \tag{6.37}$$

where $u_1, u_2 \in D$ *are given by* (6.33). *Assume* $\hat{w}_a = (w_{a1}, w_{a2}) \in \hat{D}$ *and* $\hat{w}_b = (w_{b1}, w_{b2}) \in \hat{D}$ *satisfy* (6.35). *In addition, it will be assumed that* $\hat{w}_a$ *satisfies* (6.36*a*), *while* $\hat{w}_b$ *satisfies* (6.36b). *Then*:

(*i*) u_1 *and* u_2 *satisfy the initial and terminal value problems respectively, if and only if* w_1 *and* w_2 *satisfy equations* (6.36).

(*ii*)
$$2\{\Omega(\hat{w}_b)-\Omega(\hat{w}_a)\} = m_1 \int_G \left[\left(\frac{\partial(u_{a1}-u_{b1})}{\partial t}\right)^2 + \frac{\partial(u_{a1}-u_{b1})}{\partial x_i}\frac{\partial(u_{a1}-u_{b1})}{\partial x_i}\right]_{t=T}$$
$$m_2 \int_G \left[\left(\frac{\partial(u_{a2}-u_{b2})}{\partial t}\right)^2 + \frac{\partial(u_{a2}-u_{b2})}{\partial x_i}\frac{\partial(u_{a2}-u_{b2})}{\partial x_i}\right]_{t} dx \geqslant 0.$$

(*iii*)
$$\Omega(\hat{w}_a) = \Omega(\hat{w}_b)$$

if and only if

$$\frac{\partial u_{a1}}{\partial t} = \frac{\partial u_{b1}}{\partial t}, \qquad \frac{\partial u_{a1}}{\partial x_i} = \frac{\partial u_{b1}}{\partial x_i} \quad \text{on } \bar{G} \text{ at } t = T \tag{6.40a}$$

and simultaneously

$$\frac{\partial u_{a2}}{\partial t} = \frac{\partial u_{b2}}{\partial t}, \qquad \frac{\partial u_{a2}}{\partial x_i} = \frac{\partial u_{b2}}{\partial x_i} \quad \text{on } \bar{G} \text{ at } t = 0. \tag{6.40b}$$

(*iv*) *If solutions* $U_1, U_2 \in D$ *of the initial and terminal value problems respectively, exist, then*

$$2\{\Omega(\hat{w}_b)-\Omega(\hat{w}_a)\} = m_1 \int_G \left[\left(\frac{\partial(u_{a1}-U_1)}{\partial t}\right)^2 + \left(\frac{\partial(u_{b1}-U_1)}{\partial t}\right)^2 + \right.$$
$$\left.\frac{\partial(u_{a1}-U_1)}{\partial x_i}\frac{\partial(u_{a1}-U_1)}{\partial x_i} + \frac{\partial(u_{b1}-U_1)}{\partial x_i}\frac{\partial(u_{b1}-U_1)}{\partial x_i}\right]_{t=T} dx +$$
$$m_2 \int_G \left[\left(\frac{\partial(u_{a2}-U_2)}{\partial t}\right)^2 + \left(\frac{\partial(u_{b2}-U_2)}{\partial t}\right)^2 + \right.$$
$$\left.\frac{\partial(u_{a2}-U_2)}{\partial x_i}\frac{\partial(u_{a2}-U_2)}{\partial x_i} + \frac{\partial(u_{b2}-U_2)}{\partial x_i}\frac{\partial(u_{b2}-U_2)}{\partial x_i}\right]_{t=0} dx.$$

Proof. This theorem is a straightforward application of Theorem 4.3.

Many extremum principles can be derived as corollaries of Theorem 6.5. In many applications interest is centred in the initial value problem. In this case one can profit from the arbitrariness of g. We do not intend to be exhaustive in this respect. However, the following theorem is given as an example.

THEOREM 6.6. *Define the functional* $\pi\colon D \to R$ *for every* $p \in D$, *by*:

$$\pi(p) = \int_G \left[\left(\frac{\partial p}{\partial t}\right)^2 + \frac{\partial p}{\partial x_i}\frac{\partial p}{\partial x_i}\right]_{t=T} dx - 2\left\{\int_0^T \left(\int_G f_G \frac{\partial p}{\partial t} dx + \int_{S1} f_{S1} \frac{\partial p}{\partial n} dx + \right.\right.$$
$$\left.\left.\int_{S2} f_{S2} \frac{\partial p}{\partial t} dx\right) dt + \int_G \left[f_t \frac{\partial p}{\partial t} + \frac{\partial f_0}{\partial x_i}\frac{\partial p}{\partial x_i}\right]_{t=0} dx\right\}.$$

Define $D_\pi \subset D$ by the condition $p \in D_\pi$ if and only if

$$\frac{\partial^2 p}{\partial t^2} - \frac{\partial^2 p}{\partial x_i\, \partial x_i} = 0 \quad on\ \bar{G}\times[0, T], \tag{6.43a}$$

$$\frac{\partial p}{\partial t} = 0 \quad on\ S_1\times[0, T], \tag{6.43b}$$

$$\frac{\partial p}{\partial n} = 0 \quad on\ S_2\times[0, T]. \tag{6.43c}$$

Assume a solution $p_M \in D$ of the initial value problem for the wave equation exists. Then:
(i) for each $p \in D_\pi$

$$\pi(p) \geqslant \pi(p_M), \tag{6.44}$$

(ii) in addition the equality sign holds at some $p \in D_\pi$ if and only if

$$\frac{\partial p}{\partial t}(\mathbf{x}, t) = \frac{\partial p_M}{\partial t}(\mathbf{x}, t) \qquad \frac{\partial p}{\partial x_i}(\mathbf{x}, T) = \frac{\partial p_M}{\partial x_i}(\mathbf{x}, T). \tag{6.45}$$

Proof. Choose the functions g so that the right hand members in (6.36a) vanish, while $g_t \equiv g_T \equiv 0$. For functions $\hat{w} \in \hat{D}$ satisfying equations (6.35) and (6.36a):

$$\Omega(\hat{w}) = \langle Pu_1, u_2\rangle - \langle g, u_1\rangle - \langle f, u_2\rangle = \frac{\alpha}{\beta}\langle Pu_1, u_1\rangle - \frac{1}{\beta}\langle f, w_1\rangle \tag{6.46}$$

because in this case equation (4.15a) is satisfied.

At the same time (6.35b) implies

$$u_1 = \frac{1}{\alpha} w_1 \text{ at } t = T.$$

Using the initial conditions (6.35a) gives

$$\Omega(\hat{w}) = \frac{\alpha}{2\beta}\int_G \left\{ f_t^2 + \frac{\partial f_0}{\partial x_i}\frac{\partial f_0}{\partial x_i} + \left[\left(\frac{\partial p}{\partial t}\right)^2 + \frac{\partial p}{\partial x_i}\frac{\partial p}{\partial x_i}\right]_{t=T} \right\} d\mathbf{x} -$$
$$\frac{\alpha}{\beta}\left\{ \int_G \left[f_t \frac{\partial p}{\partial t} + \frac{\partial f_0}{\partial x_i}\frac{\partial p}{\partial x_i}\right]_{t=0} d\mathbf{x} + \right.$$
$$\left. \int_0^T \left(\int_G f_G \frac{\partial p}{\partial t} d\mathbf{x} + \int_{S1} f_{S1}\frac{\partial p}{\partial n} d\mathbf{x} + \int_{S2} f_{S2}\frac{\partial p}{\partial t} d\mathbf{x}\right) dt \right\} \tag{6.48}$$

where αp stands for w_1. Multiply by $2\beta/\alpha$ and eliminate an irrelevant constant to obtain the functional $\pi(p)$ given by (6.42).

This research was supported by the Instituto de Investigaciones en Matematicas Aplicadas y en Sistemas, National University of Mexico (I.H.), the Department of Mathematics, University of Reading, U.K. (M.J.S.), and the United States Army under Contract No. DAAG29-75-C-0024 (M.J.S.).

References

ARTHURS, A. M. 1970 *Complementary Variational Principles.* Oxford University Press.
BREZIS, H. & EKELAND, I. 1976 *C.R. Acad. Sci., Paris* **A282**, 971–974.
COLLINS, W. D. 1976 *Technical Summary Report 1624.* Mathematics Research Center, University of Wisconsin.

HERRERA, I. 1974 *A general formulation of variational principles.* Instituto de Ingeniería, UNAM, E-10.
HERRERA, I. & BIELAK, J. 1974 *Archs ration. Mech. Analysis* **53**, 131–149.
HERRERA, I. & BIELAK, J. 1976 *Q. appl. Maths* **34**, 85–102.
MORSE, P. M. & FESHBACH, H. 1953 *Method of Theoretical Physics.* Volume I. McGraw-Hill.
NOBLE, B. & SEWELL, M. J. 1972 *J. Inst. Maths Applics* **9**, 123–193.
ROBINSON, P. D. 1971 In *Nonlinear Functional Analysis and Applications* (Ed. L. B. Rall). Academic Press, pp. 507–576.
SEWELL, M. J. 1969 *Phil. Trans. R. Soc.* **A265**, 319–351.
SEWELL, M. J. 1973*a* *J. Struct. Mech.* **2**, 1–32.
SEWELL, M. J. 1973*b* *J. Struct. Mech.* **2**, 135–158.
SEWELL, M. J. & NOBLE, B. 1976 *Technical Summary Report* 1703. Mathematics Research Center, University of Wisconsin.