Int. J. Engng Sci. Vol. 22, No. 11/12, pp. 1343–1371, 1984
Printed in the U.S.A.

0020–7225/84 $3.00 + .00
Pergamon Press Ltd.

VARIATIONAL FORMULATION FOR EVERY NONLINEAR PROBLEM

E. TONTI
Istituto di Scienza delle Costruzioni, Universitá di Trieste, Piazzale Europa 1, Trieste, Italia

Abstract—It is shown that for every linear or nonlinear problem, whose solution exists and is unique, one may find many functionals whose minimum is the solution of the problem. They are obtained after a transformation of the given problem into another by the application of an "integrating operator": this transforms a differential problem into an integro-differential one. The procedure used to obtain such functionals is straightforward and is described in detail. Examples are exhibited and the numerical effectiveness of the method is tested. The variational formulation so obtained contains the classical formulation as a particular case when it exists.

1. INTRODUCTION

THE PROBLEM of giving a variational formulation to all linear or nonlinear problems has so far remained unsolved.

There is a precise test to find out whether a given problem has a variational formulation or not. This test is the symmetry of the operator, if this is linear, or the symmetry of its Gateaux derivative, if the operator is nonlinear.

The essential point, no usually stressed, is that the symmetry of a linear operator, like that of a matrix, is not an absolute notion: it is relative to a bilinear form.

If the given problem does not meet this symmetry condition, one may try to transform the given problem into another one that meets the symmetry condition.

Three methods can be used:

1. To transform the given problem into another one with the same solutions.
2. To change the bilinear form (which is implicit in a variational formulation).
3. To change the function.

As we shall see, the first two methods are equivalent (Section 3) so that every transformation of the original problem corresponds to a bilinear form and vice versa.

There are equations that have long resisted a variational formulation. This is the case of the Fourier equation (heat conduction) and of the Fick equation (diffusion), which was given a truly variational formulation around 1964. This result was obtained by Gurtin[14], who showed how to give a variational formulation to linear initial value problems having an equation with constant coefficients.

Gurtin's idea was to make a preliminary transformation of an equation into an integro-differential one and the introduction of the convolution product of two functions. This method has led to the variational formulation of many linear initial value problems and a great number of papers which have appeared mainly in engineering reviews.

Gurtin's method was simplified in 1972 by the author[46] who showed that one may introduce a convolutive bilinear form to give a variational formulation to linear initial value problems having an equation with constant coefficients.

The idea of adapting the bilinear form to the given operator was perfected by Magri in 1974[24], he showed that every linear equation admits of a variational formulation, giving an explicit method for obtaining the functional.

This result contradicts the common belief that equations admitting a variational formulation form a priviledged class. At this stage every linear problem may be associated with many functionals; for each of these the stationary value is attained in correspondence to the solution to the problem. In general, these functionals are not extreme at this point.

In 1978 Reiss and Haug [34], using Magri's results, explored the possibility of finding among the many functionals those that give an extremum principle for linear initial value problems.

But what about the larger class of nonlinear problems?

Apart from the particular results obtained by the trial-and-error method applied to single equations, a first attempt at an extension of Magri's result was made by Telega in 1979[39] (see Appendix 5). The method we have followed may be summarized as follows. To give a general procedure for transforming the given problem into another problem with the same solution but such that it admits of a variational formulation. This idea evolved from the theory of the integrating factor. By "integrating factor" is usually meant a function or, more generally, a matrix whose entries are numbers of functions. In all cases an essential requirement is that the factor be invertible to assure that no new solutions are added to the given problem.

As an example let us consider the boundary value problem

$$\begin{cases} m\ddot{q} + h\dot{q} + kq = f & 0 \leqslant t \leqslant T \\ q(0) = 0 \quad q(T) = 0 & q(t) \in C^2(0, T). \end{cases} \tag{1.1}$$

Since the operator is not symmetric, the problem does not admit of a variational formulation. If we multiply the equation by the integrating factor

$$p(t) = \exp(ht/m) \tag{1.2}$$

the problem becomes

$$m\frac{\mathrm{d}}{\mathrm{d}t}\left[\exp(ht/m)\frac{\mathrm{d}q}{\mathrm{d}t}\right] + k\exp(ht/m)q = \exp(ht/m)f \tag{1.3}$$

which is symmetric. Now a variational formulation is possible. The functional is

$$F[q] = \int_0^T \exp(ht/m)\left[\frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2 - fq\right]\mathrm{d}t. \tag{1.4}$$

In general the existence of the integration factor for the problem

$$f(t; q, \dot{q}, \ddot{q}) = 0 \tag{1.5}$$

is put in the following terms: *does a function* $r(t; q, \dot{q})$ *exist such that the equation*

$$r(t; q, \dot{q})f(t; q, \dot{q}, \ddot{q}) = 0 \tag{1.6}$$

is the Euler–Lagrange equation of a functional? An obvious requirement is that $r(t, q, \dot{q})$ should not vanish identically for some function $q(t)$.

The problem stated in this form seldom admits a solution. The function $r(t; q, \dot{q})$ must satisfy a partial differential equation that expresses the integrability condition. An explicit solution of this equation in general is not possible: particular solutions can be found by trial-and-error.

But why should we limit our search to an integrating *factor* and not consider an integrating *operator?* When we differentiate both sides of an equation or perform a Laplace transform or multiply a system of differential equations by a matrix, we apply an *operator.*

Then one may search for an integrating operator R such that the equation

$$R(q; f(t; q, \dot{q}, \ddot{q})) = 0 \tag{1.7}$$

has the same solution as the old one. If one puts the problem in this form one can prove that *not only does an integrating operator always exist but that there is an infinite number of such operators.* These integrating operators can be found explicitly, as can the functionals.

It turns out that the new equation has the form

$$\int_0^T r(t, q(t), \dot{q}(t); \tau, q(\tau), \dot{q}(\tau)) f(\tau, q(\tau), \dot{q}(\tau), \ddot{q}(\tau))\, d\tau = 0, \tag{1.8}$$

i.e. the transformed equation becomes integro-differential.

Moreover, integrating operators can be found such that the functional is an extremum at the solution and so that direct methods of the Calculus of Variations may be applied. This statement is valid for linear or nonlinear equations, with total or partial derivatives, of any order (odd or even). The equation may be of an integral or integro-differential type, or it may be an equation with retarded argument, or it may even be a system of differential or integral equations, etc.

In this way we may obtain a variational formulation for nonlinear initial value problems that, up to now, have been excluded from a variational formulation.

In doing so, we lose the possible physical meaning of the functional; this is counter-balanced by the fact that the minimization of the functional can be used for numerical calculations. This is what we show in Section 7.

2. NOTATION

We shall use an operational approach because it is both synthetic and conceptually simple. To make it technically simple as well, we shall clarify what is understood by the operational notation.

In talking about equations, we mean any kind of equation; whether differential, integral, or integro-differential; both a single equation and a system of equations and linear or nonlinear. This paper deals essentially with nonlinear operators: when the operator is linear this fact will be expressly stated.

An equation is usually associated with additional conditions that specify initial, boundary and regularity conditions, as well as the functional class. The functions considered may be real or complex-valued, scalar vector or tensor-valued. The set formed of an equation and all additional conditions constitutes a *problem*. Every problem may be written in the general form

$$N(u) = \emptyset_V \tag{2.1}$$

where N denotes an operator and u is a function or a set of functions. This notation is analogous to that of matrix theory, such as $Lu = 0$ or $L\{u\} = \{0\}$. The round brackets in (2.1) are customary in the theory of nonlinear operators, just as the notation $f(x) = 0$ is customary in the theory of functions.

The set of elements u that satisfies the prescribed initial or boundary conditions and the given functional class is called the *domain* of the operator and will be denoted as $D(N)$. This domain will be conceived of as a subset of a vector space U.

The set of elements $v = N(u)$ constitutes the *range* of the operator and will be denoted by $R(N)$; it is supposed to be embedded in another vector space V. $\emptyset_V$ denotes the null element of the V-space. The two spaces U and V may coincide: they are not considered to be Banach or Hilbert spaces: a norm will be introduced after the introduction of a bilinear functional.

A linear problem written in the form

$$Lu = f \tag{2.2}$$

where L: $D(L) \subset U \to V$ denotes a linear operator, may be included in the general form (2.1) by letting $N(u) = Lu - f$. This operator is nonlinear; in particular it is an affine operator.

The notations $N(u) = \emptyset_V$ or $Lu = f$ do not rule out that there may be functions,

different from u and f, that make up the operator. So the problem

$$\begin{cases} \mathrm{d}/\mathrm{d}x[p(x)\,\mathrm{d}u(x)/\mathrm{d}x] + q(x)u(x) = f(x) \\ u(0) = 0 \quad u(1) = 0 \end{cases} \tag{2.3}$$

may be written

$$L(p(x), q(x); u(x)) = f(x) \tag{2.4}$$

or simply

$$Lu(x) = f(x). \tag{2.5}$$

In an analogous way an equation that links a "source" function f with a "configuration" function u may be written

$$N(f; u) = \emptyset_V \tag{2.6}$$

or simply

$$N(u) = \emptyset_V \tag{2.7}$$

depending on whether we want to stress the presence of an assigned function $f(t)$ or not.

We may remark that, in the case of differential operators, the linearity of the operator implies both the linearity of the formal differential operator and the linearity of its domain, which in turn requires homogeneous boundary or initial conditions.

Non-homogeneous linear boundary conditions make the domain a convex set and render the operator nonlinear (in particular affine).

Let us perform the following decomposition

$$N(u + \epsilon w) - N(u) = L(u; \epsilon w) + R(u; \epsilon w) \tag{2.8}$$

where w is such that $u + \epsilon w$ belongs to the domain of N for every u.

L is a linear operator on w that contains u in an arbitrary way; R is a nonlinear operator on w also containing u such that

$$\lim_{\epsilon \to 0} R(u; \epsilon w)/\epsilon = \emptyset_V \quad \text{for every} \quad u \in D(N). \tag{2.9}$$

The limit is defined by the topology of the V-space. The element $v = L(u; \epsilon w)$ of the V-space is called the *Gateaux differential* of N at u and the operator $L = L(u; \,.\,)$ is called the *Gateaux derivative* of N at u.†

One may write $L(u; \,.\,) = N'_u(u; \,.\,)$. From eqn (2.8) and taking into account eqn (2.9) we have

$$N'_u(u; w) = N'_u w = \mathrm{d}/\mathrm{d}\epsilon[N(u + \epsilon w)]_{\epsilon = 0}. \tag{2.10}$$

The second notation makes clear the linearity on w (no brackets), while the first one points up the dependence on u. In other words, the operator $N'_u(u; w)$ is an operator linear on w and has variable coefficients that are formed of u and its derivatives. Letting $\delta u = \epsilon w$, the linear part of the variation (Gateaux differential) may be denoted by $\delta N(u)$. Then

$$\delta N(u) = N'_u(u; \delta u) = N'_u u. \tag{2.11}$$

† In the author's paper[47] the Gateaux derivative was erroneously called the Frechét derivative.

The symbol δ applied to operators is an extension of a similar symbol used for functionals in the Calculus of Variations. If F denotes a functional, one may write

$$\delta F[u] = F'_u[u; \delta u] = F'_u \delta u. \tag{2.12}$$

To give a variational formulation to a problem we need a bilinear functional. The reason for this may be seen by observing that there is a surprising analogy between the notion of an operator, i.e. a mapping between two function spaces U and V, and that of a vector field in 3-dimensional space. A vector field may be defined as a mapping of points of the 3-dimensional space on vectors of another 3-dimensional space. In this analogy, the point of the 3-dimensional space corresponds to the point of a function space, i.e. a function; a line in 3-space corresponds to a one parameter family in function space. The notion of a scalar product between a vector $\mathbf{v}$ and an infinitesimal vector $\mathbf{dr}$ in 3-dimensional space corresponds to a bilinear functional $\langle v, \delta u \rangle$ where $\delta u \in U$ and $v \in V$.

Now the question of knowing whether a given problem admits of a variational formulation is similar to the question of knowing if a vector field admits of a potential. In the latter case the necessary condition is that the circulation $v(\mathbf{r}) \cdot \mathbf{dr}$ along a line be dependent only on the extreme points. In the former case, the analogous condition is that the "circulation" $\langle N(u), \delta u \rangle$ must be dependent only on the extreme functions of the "line" in function space. The reader may find an elementary exposition of these notions in [47].

A *bilinear functional* that we shall denote as $\langle v, u \rangle$, is a map B: $V \times U \to R$ that must fulfil the following requirements.

(1) It must be *real-valued* (even if U and V are vector spaces over the complex number field).

(2) It must be *bilinear* over the real number field.

(3) It must be *non-degenerate*, i.e.

$$\text{if} \quad \langle v, u_0 \rangle = 0 \quad \text{for every} \quad v \in V \quad \text{then} \quad u_0 = \phi_v$$

$$\text{if} \quad \langle v_0, u \rangle = 0 \quad \text{for every} \quad u \in U \quad \text{then} \quad v_0 = \phi_v.$$

The real number $s = \langle v, u \rangle$ is then called the *scalar product* of the two elements $v \in V$ and $u \in U$; the V-space is called the *dual* of the U-space and one writes $V = U^*$. One also says that U and V are put into duality by the bilinear functional $\langle v, u \rangle$, which is called "canonical" to distinguish it from other possible bilinear functionals.

The bilinear functionals used in practice are of the following kinds. Let us denote as $\mathbf{v} \cdot \mathbf{u}$ the local scalar product, i.e. a scalar valued function formed of two vector-valued or two tensor-valued functions of opposite variance, such as

$$\sum_k v_k(\mathbf{x}) u^k(\mathbf{x}) \qquad \sum_{hk} v_{hk}(\mathbf{x}) u^{hk}(\mathbf{x}). \tag{2.13}$$

In the case of tensors of second rank, the two tensors must have the same symmetry: both must be symmetrical, as in the case of continum mechanics with stress and strain tensors, or both must be anti-symmetrical, as in the case of the electromagnetic tensor $F_{\alpha\beta} G^{\alpha\beta}$.

If we denote by Ω a subset of R^n and by $\mathbf{x}$ a point of Ω, i.e. $\mathbf{x} = (x^1, x^2, \ldots x^n)$, we shall take as "canonical" the bilinear functional

$$\langle v, u \rangle = \int \mathbf{v}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) \, d\Omega. \tag{2.14}$$

This is the bilinear functional usually employed in physical theories. It is not the most general bilinear functional. If $A : U \to U$ and $B : V \to V$ are two linear invertible operators whose domains are the entire U- and V-spaces respectively, a more general non-degenerate

bilinear functional is

$$\langle v, u \rangle = \int B\mathbf{v}(\mathbf{x}) \cdot A\mathbf{u}(\mathbf{x}) \, d\Omega. \tag{2.15}$$

The requirement of non-degeneracy is a restriction both on the pairs of spaces U and V and on the mappings N that are candidates for a variational formulation. For example a problem such as

$$\begin{cases} \operatorname{div} \mathbf{D}(\mathbf{x}) = \rho(x) \\ \mathbf{n} \cdot \mathbf{D}|_{S_1} = 0 \end{cases} \tag{2.16}$$

cannot as it stands be a candidate for a variational formulation, because on the spaces of scalar-valued functions $\rho(x)$ and those of vector-valued functions $\mathbf{D}(\mathbf{x})$ a non-degenerate bilinear functional cannot be defined.

Of course, to problem (2.16) the corresponding adjoint problem may be added

$$\begin{cases} \mathbf{D}(\mathbf{x}) = -\epsilon \operatorname{grad} V(\mathbf{x}) \quad \text{with} \quad S_1 u S_2 = \delta\Omega. \\ V(\mathbf{x})|_{S_2} = 0 \end{cases} \tag{2.17}$$

The whole problem made up of (2.16) and (2.17) becomes a possible candidate for a variational formulation because on the spaces of scalar valued functions $\rho(x)$ and $V(x)$ one may define a non-degenerate bilinear functional.

Note—the two spaces U and V are required to be linear over the real or complex number field. To make a variational formulation one needs a real-valued, non-degenerate bilinear functional over the *real* number field.

In dealing with complex-valued functions, one has to deal with *hermitean* functionals (in $U \times U$) or *sesquilinear* functionals (in $V \times U$): these are complex-valued non-degenerate functionals that fulfil the requirements

$$\langle \lambda v \,|\, u \rangle = \lambda \langle v \,|\, u \rangle \quad \langle v \,|\, \lambda u \rangle = \bar{\lambda} \langle v \,|\, u \rangle \tag{2.18}$$

where λ is an arbitrary complex number and $\bar{\lambda}$ its complex conjugate. These functionals are linear with respect to the first element and "half-linear" for the second one. To deal with the variational formulation, one must construct the functional on $V \times U$

$$\langle v, u \rangle = \frac{1}{2}[\langle v \,|\, u \rangle + \overline{\langle v \,|\, u \rangle}] \tag{2.19}$$

which is real-valued, non-degenerate and bilinear on the real number field. Every operator that is hermitean with respect to the sesquilinear functional (2.18) is automatically symmetric with respect to the bilinear functional (2.19).

Once a bilinear functional has been introduced it is natural to introduce a topology in both spaces in such a way that the bilinear functional will be continuous with respect to both arguments. Such a topology is said to be compatible with duality.

For example, if we have an operator N whose domain is that of differentiable functions defined in (0, 1) that vanish at $x = 0$, and if the bilinear functional is the usual one (2.14), one may choose $D(N) = U = C^1(0, 1)$ and $R(N) = V = C(0, 1)$. The topologies induced by the norms

$$\|u\| = \max_{x \in (0,1)} [|u(x)| + |u'(x)|] \qquad \|v\| = \max_{x \in (0,1)} |v(x)| \tag{2.20}$$

are compatible with duality. To prove this let us observe that the two spaces are complete and if $|u_n - u_0| \leqslant \epsilon$

$$\begin{aligned} |\langle v, u_n \rangle - \langle v, u_0 \rangle| = |\langle v, u_n - u_0 \rangle| &= \left| \int_0^1 v(u_n - u_0) \, dx \right| \\ &\leqslant \int_0^1 |v(u_n - u_0)| \, dx = \int_0^1 |v| \, |u_n - u_0| \, dx \leqslant (\max |v|) \cdot \epsilon \cdot 1. \end{aligned} \tag{2.21}$$

Since v is continuous and the interval is finite the property follows. Once the continuity is assured, if u belongs to a dense subset D of a linear space U the condition $\langle v_0, u \rangle = 0$ for every $u \in D$ assures that v_0 is the null element of V.

3. VARIATIONAL FORMULATION

Let F be a functional, i.e. a mapping $D(F) \subset U \to R$. If we consider a variation δu we can write the corresponding variation δF as

$$\delta F[u] = \langle E(u), \delta u \rangle \tag{3.1}$$

as is usual in the "δ-process" of the Calculus of Variations.

The relation between E and F can be expressed by saying that the operator E: $D(E) \subset U \to V = U^*$ is the *gradient* of F, and in turn that the functional F is the *potential* of E. An operator E that is the gradient of a functional is called a *potential operator*. An element u_0 such that $\delta F|_{u_0} = 0$ for any $u_0 \subset U$ is called a *critical point* of the functional F. This is said to be *stationary* at u_0.

The element $v = E(u)$ is called the *variational derivative* of the functional F, or even its *functional derivative* [30]. When E is a differential operator, if we denote by $\mathscr{E}$ the corresponding formal differential operator, the equation

$$\mathscr{E}(u) = 0 \tag{3.2}$$

is the usual *Euler equation* while

$$E(u) = \emptyset_V \tag{3.3}$$

is the *Euler problem*. We may remark that a variational formulation makes sense for a problem, but not for an equation: both boundary and initial conditions are essential ingredients of a variational formulation.

The problem of finding out whether a given operator is the gradient of a functional and of finding the functional is known as the *inverse problem* of the Calculus of Variations. Conversely, the simple problem of finding the gradient of a given functional may be called the *direct problem*.

We now have the concepts needed to distinguish between two kinds of variational formulation:

(a) *Variational formulation in the restricted sense*

Given a problem $N(u) = \emptyset_V$ with N: $D(N) \subset U \to V = U^*$, find a functional F, if any, such that the operator N is the gradient of F, i.e. such that

$$\delta F = \langle N(u), \delta u \rangle. \tag{3.4}$$

This implies that the solutions to the problem are the critical points of the functional F and vice versa.

This form of the inverse problem is the one given by Hirsch in 1897 ([19], p. 52).

(b) *Variational formulation in the extended sense*

Given a problem $N(u) = \emptyset_V$ with N: $D(N) \subset U \to V = U^*$, find a functional $\bar{F}$, if any, whose critical points are solutions to the problem and vice versa.

This implies that for a given operator N an operator $\bar{N}$ exists such that

$$\delta \bar{F} = \langle \bar{N}(u), \delta u \rangle \tag{3.5}$$

and the problems $N(u) = \emptyset_V$ and $\bar{N}(u) = \emptyset_V$ have the same solutions.

This statement is less demanding than the former one, because it requires only that the

critical points should coincide with the solutions, without positing the additional requirement that N must be the gradient of the functional. The gradient of the functional $\bar{F}$ will be another operator, say $\bar{N}$, that will be linked in some way to the operator N.

Figure 1 shows the relation between the two formulations.

This form of the inverse problem is the one given by Davis in 1928[5] and used by Douglas in 1941 ([9], p. 71).

The existence of the variational formulation in the restricted sense for a given problem is based on the following fundamental theorem.

Theorem 1 (Volterra, 1913 [52])

In order that an operator N: $D(N) \subset U \to R(N) \subset V = U^*$ *be the gradient of a functional it is necessary that the circulation of the element* $v = N(u)$ *along any reducible closed line contained in* $D(N)$ *vanishes. Taking an infinitesimal parallelogram the vanishing of the circulation is expressed by*

$$\langle N'_u \varphi, \psi \rangle = \langle N'_u \psi, \varphi \rangle, \tag{3.6}$$

i.e. the operator $N'_u(u; .)$ *must be symmetric.*

If the domain $D(N)$ *is simply connected, condition* (3.6) *becomes sufficient. In this case if* $\eta(\lambda)$ *denotes a one-parameter family of elements (a "line" from* u_0 *to* u*) with* $\eta(0) = u_0$ *and* $\eta(1) = u$*, the functional is given by*

$$F[u] = F[u_0] + \int_{\lambda=0}^{\lambda=1} \left\langle N(\eta(\lambda)), \frac{\partial \eta}{\partial \lambda} \right\rangle \mathrm{d}\lambda. \tag{3.7}$$

A historical remark

While the symmetry condition (3.6) was established by Volterra in 1887 ([53], p. 104), the preceding theorem was established for the first time by Volterra in 1913 ([52], p. 47), where he wrote:

"On peut tirer de même de la formule (12) la condition pour que

$$\int_a^b X\,|\,x_a^b(t), \eta]\,|\,\delta x(\eta)\,\mathrm{d}\eta$$

soit la différentiale totale exacte d'une fonction de ligne que nous saurons alors calculer. Il faut que

$$X'\,|\,[x_a^b(t), \eta, \xi]\,| = X'\,|\,[x_a^b(t), \xi, \eta]\,|,$$

le second paramètre indiquant toujours dans les expressions precédéntes, le point où l'on effectue la dérivation.

C'est la *condition de symmétrie de la dérivée seconde* que nous avions déjà indiquée.

On peut énconcer le résultat précis suivant. *La condition nécessaire et suffisante pour*

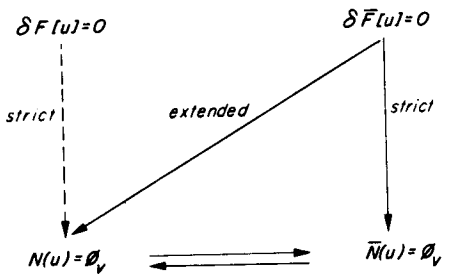


Fig. 1. The restricted and extended variational formulation.

que

$$X|[x_a^b(t), \eta]|$$

soit la dérivée d'une fonction de ligne

$$F|[x_a^b(t)]|$$

est que

$$X'|[x_a^b(t), \eta, \xi]| = X'|[x_a^b(t), \xi, \eta]|.$$

On saura calculer cette fonction de ligne.

On aura

$$F|[x(t)]| - F|[x_0(t)]| = \int_{S_1}^{S_2} \mathrm{d}s \int_a^b X|[x(t_a^b|s), \eta]| \frac{\mathrm{d}x(\eta/s)}{\mathrm{d}s} \mathrm{d}\eta."$$

Furthermore, we observe that in 1918 Evans published a book in which he wrote ([10], p. 23]: "The condition

$$\phi''[C|MM_1] = \phi''[C|M_1M]$$

. . .was originally stated by Volterra."

In 1933 Kerner ([21], p. 550) reported these results and quoted Volterra (his Ref. [4]). In 1964 Vainberg ([50], p. 56) gave the same results, quoting Kerner (his Ref. [42b]). In 1969 Tonti [47] popularized these results, quoting Vainberg (his Ref. [1]). The result of all this is that some authors attribute this theorem to Kerner see [30], [32], p. 42), ([39], p. 176); others to Vainberg (see [3], p. 33), ([33], p. 75), ([51], p. 1179) and still others to the present author (see [38]).

The fundamental observation is that the symmetry condition involves the use of a bilinear functional. This implies that if an operator does not satisfy the symmetry condition (3.6), one may change the bilinear functional. This is one key for giving a variational formulation to problems that, when classical bilinear functionals are used, do not admit of one.

We may summarize this fact in the following sentence. *The variational formulation requires symmetry, and the symmetry is related to a bilinear functional.*

Let us take an example. The operator D given by

$$D = \{\mathrm{d}/\mathrm{d}t, u(0) = 0, u \in C^1(0, T)\} \tag{3.8}$$

which is not symmetric with respect to the usual bilinear functional

$$\langle v, u \rangle = \int_0^T v(t)u(t)\,\mathrm{d}t \tag{3.9}$$

becomes symmetrical with respect to the convolutive bilinear functional

$$\langle v, u \rangle_c = \int_0^T v(T-t)u(t)\,\mathrm{d}t = \int_0^T v(t)u(T-t)\,\mathrm{d}t \tag{3.10}$$

as is easily shown [46].

If C denotes the convolution operator defined by

$$Cv(t) = v(T-t) \tag{3.11}$$

we may write

$$\langle v, u\rangle_c = \langle Cv, u\rangle = \langle v, Cu\rangle \tag{3.12}$$

because C is a symmetrical operator.

The symmetry of D with respect to the convolutive functional, i.e.

$$\langle Du, u'\rangle_c = \langle Du', u\rangle_c \tag{3.13}$$

is equivalent to the statement that CD is symmetric with respect to the canonical bilinear functional

$$\langle CDu, u'\rangle = \langle Du, u'\rangle_c = \langle Du', u\rangle_c = \langle CDu', u\rangle. \tag{3.14}$$

This is a general fact: *the change in bilinear functional is equivalent to pre-multiplication by an operator.*

This result implies that to say that the symmetry of an operator is related to a bilinear functional is equivalent to saying that the operator may be made symmetric by application of an integrating operator.

Thus, instead of changing the bilinear functional we may search for an integrating operator. This is a second key to use in trying to give a variational formulation.

For linear operators, the choice of one or the other point of view is a matter of taste, but for nonlinear operators it is simpler to use the canonical bilinear functional and apply an integrating operator (otherwise the bilinear functional would depend on u, i.e. using the language of differential geometry, it would be a local bilinear functional).

It is easily seen that in this equivalence the requirement that the new bilinear functional be non-degenerate is equivalent to the requirement that the integrating operator be invertible (i.e. kernel-free).

4. EXTENDED VARIATIONAL FORMULATION

In exploring, from a general point of view, the possibility of finding an integrating operator for a given operator we shall take an induction approach.

In matrix theory a system $Au = b$ may always be transformed into another system by pre-multiplication by a matrix C. If C is invertible, the new system $CAu = Cb$ has the same solutions as the old one. In particular, this is the case when the adjoint matrix A^* is invertible: the system becomes $A^*Au = A^*b$. In this case the vector u that solves the system makes the function stationary

$$f(u) = \|Au - b\|^2 \tag{4.1}$$

and this gives rise to the least square method.

Thus, if A is a square matrix and if A^* is invertible, the solution of the original system $Au = b$ makes the function (4.1) stationary.

In general, the same procedure cannot be performed on differential operators. To show why, let us consider the linear differential operator

$$D = \left\{\frac{\mathrm{d}}{\mathrm{d}t},\quad u(0) = 0,\quad u \in C^1(0, T)\right\} \tag{4.2}$$

and the problem

$$Du = f \quad \text{with} \quad f \in C(0, T). \tag{4.3}$$

The adjoint operator is

$$D^* = \left\{-\frac{\mathrm{d}}{\mathrm{d}t}, v(T) = 0, v \in AC(0, T)\right\} \tag{4.4}$$

where $AC(0, T)$ denotes the class of absolutely continuous functions. It is true that D^*D is a symmetric operator, but D^* cannot be applied to both the members of the problem (4.3) because f does not fulfil the final condition $f(T) = 0$ and thus does not belong to the domain of definition of D^*. When the method is applied to differential equations, it is assumed that $f \in D(A)$ ([25], p. 496).

In other words, the domain of the operator D^*D is a restriction on the domain of D. If $Du = f$ is an equation representing a physical law, all elements $f \in R(D)$ describe possible sources. Thus, the restriction of the domain required by the application of the operator D^* excludes elements f and, therefore, possible source distributions: this cannot be accepted.

We may think of applying to the problem (4.3) an integral operator, K: $R(N) \rightarrow U$, i.e. we may try to perform an integral transformation, so that f will be transformed into a new function $\bar{f}$ that fulfils the final condition $\bar{f}(T) = 0$. We may choose an operator like the following

$$\bar{f}(t) = \int_0^T k(t, \tau) f(\tau)\, d\tau \tag{4.5}$$

with the condition $k(T, \tau) = 0$. Or we may choose an integral operator of the kind of Volterra

$$\bar{f}(t) = \int_0^t k(t, \tau) f(\tau)\, d\tau. \tag{4.6}$$

The problem (4.3) becomes

$$KDu = Kf. \tag{4.7}$$

We can now apply the operator D^*

$$D^*KDu = D^*Kf. \tag{4.8}$$

The new problem (4.8) has the same solutions as the given one if D^* and K are invertible operators. Moreover if the integral operator K is symmetrical, the operator D^*KD is also symmetrical, and thus the Volterra condition (6.1) is fulfilled. This means that problem (4.8) admits of a restricted variational formulation and thus, problem (4.2) admits of an extended variational formulation.

As we can see, the role of the integral operator K is to modify the range of D, making it "digestible" by the operator D^*. This problem does not arise in matrix theory because the domain of the matrix is the whole vector space.

The procedure indicated above may be extended to nonlinear operators, as is shown by the following theorem [40].

Theorem 2

Let us consider the problem

$$N(u) = \phi_V \tag{4.9}$$

where N is a nonlinear operator N: $D(N) \subset U \rightarrow R(N) \subset U^*$ such that (1) *the solution of the problem exists and* (2) *it is unique*; (3) *$D(N)$ is simply connected*; (4) *$N'_u(u; .)$ exists*; (5) *$D(N'_u)$ is dense in U*; (6) *$N'^*_u(u; .)$ is invertible for every $u \in D(N)$.*

Then for every operator K that fulfils the following conditions: (7) $D(K) \supset R(N)$; (8) $R(K) \subset D(N'^*_u)$; (9) *it is linear*; (10) *it is invertible*; (11) *it is symmetrical, the operator $\bar{N}$ defined by*

$$\bar{N}(u) = N'^*_u(u; KN(u)) \tag{4.10}$$

has the following properties: (*a*) *its domain coincides with that of* N; (*b*) *the problems* $N(u) = \phi_V$ *and* $\bar{N}(u) = \phi_V$ *have the same solution*; (*c*) *it is a potential operator.*

From properties (*b*) *and* (*c*) *it follows that the solution of problem* (4.9) *is the critical point of the functional*

$$\bar{F}[u] = \frac{1}{2}\langle N(u), KN(u)\rangle \tag{4.11}$$

whose gradient is the operator $\bar{N}$. *The functional vanishes when the solution is reached.*

Moreover if (12) K *is positive definite, then* (*d*) $\bar{F}[u]$ *is minimum at the critical point.*

Proof. Figure 2 will help us to simplify the proof: the operators are represented by pipes, the ingress represents the domain, the egress represents the range.

It can be easily seen that conditions (7) and (8) are represented by the fact that the ingress to the second pipe contains the egress from the first one.

The operator K is represented by the central part of the pipe. It is then clear that the operators N and $\bar{N}$ have the same domain: (a) is proved.

The linearity of N'_u and of N'^*_u, together with the linearity of K, also implies that $N'^*_u(u; K\varphi)$ is linear on φ: this assures that ϕ_V is mapped into itself. Then if u_0 is the solution to problem (4.9), it follows that

$$\bar{N}(u_0) = N'^*_{ii}(u_0; KN(u_0)) = N'^*_u(u_0; K\phi_V) = N'^*_u(u_0; \phi_V) = \phi_V. \tag{4.12}$$

This means that u_0 is also the solution to the problem $\bar{N}(u) = \phi_V$.

Conversely, let u_0 denote a solution of $\bar{N}(u) = \phi_V$. Conditions (6) and (10) assure that θ_V is mapped into itself. In fact, since

$$N'^*_u(u_0; KN(u_0)) = \phi_V \tag{4.13}$$

if we apply $(N'^*_u)^{-1}$ to both members

$$KN(u_0) = (N'^*_u)^{-1}\phi_V = \phi_V \tag{4.14}$$

and then apply K^{-1} we obtain

$$N(u_0) = K^{-1}\phi_V = \phi_V. \tag{4.15}$$

Then u_0 is also solution of the problem (4.9). Conditions (2), (6) and (10) assure that a sequence like, for instance, that represented by a dotted line in Fig. 2 cannot arise. Thus, (b) is proved.

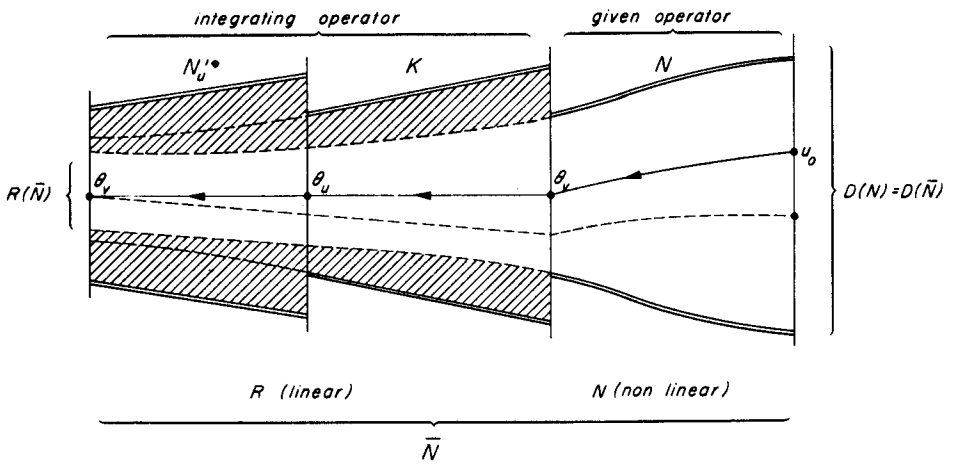


Fig. 2. Relation between the operators.

To prove the existence of the potential one may perform the test of symmetry of N'_u: if this is satisfied and since the domain is simply connected for condition (3), the existence of the potential follows. Alternatively one may find the potential directly: we shall follow the latter method, while the former is developed in Appendix 1.

We have (see (3.7))

$$\int_0^1 \left\langle \bar{N}(\eta), \frac{\partial \eta}{\partial \lambda} \right\rangle d\lambda = \int_0^1 \left\langle N'^*_\eta(\eta; KN(\eta)), \frac{\partial \eta}{\partial \lambda} \right\rangle d\lambda$$

$$= \int_0^1 \left\langle N'_\eta\left(\eta; \frac{\partial \eta}{\partial \lambda}\right), KN(\eta) \right\rangle d\lambda = \int_0^1 \left\langle \frac{\partial N(\eta)}{\partial \lambda}, KN(\eta) \right\rangle d\lambda$$

$$= \int_0^1 \frac{\partial}{\partial \lambda} \left[\frac{1}{2} \langle N(\eta), KN\langle \eta) \rangle \right] d\lambda = \left[\frac{1}{2} \langle N(\eta), KN(\eta) \rangle \right]_{\lambda=\emptyset}^{\lambda=1}$$

$$= \frac{1}{2} \langle N(u), KN(u) \rangle - \frac{1}{2} \langle N(u_0), KN(u_0) \rangle. \tag{4.16}$$

Equations (4.16) tell us that a functional exists and is given by

$$\bar{F}[u] = \frac{1}{2} \langle N(u), KN(u) \rangle. \tag{4.17}$$

Let us prove that the gradient of $\bar{F}$ is $\bar{N}$

$$\delta \bar{F}[u] = \langle N(u), \delta KN(u) \rangle = \langle N(u), K\delta N(u) \rangle = \langle N(u), KN'_u(u; \delta u) \rangle = \langle N'_u(u; \delta u),$$
$$KN(u) \rangle = \langle N'^*_u(u; KN(u)), \delta u \rangle = \langle \bar{N}(u), \delta u \rangle \tag{4.18}$$

then the vanishing of $\delta \bar{F}[u]$ implies the vanishing of the last term for every δu. Since δu belongs to the domain of N'_u, which is dense in U, and since the bilinear functional is continuous, it follows that the condition $\delta F[u] = 0$ implies $\bar{N}(u) = \emptyset_V$. This means that the critical point of the functional coincides with the solution to the problem $N(u) = \emptyset_V$ and therefore with that of the given problem.

Finally, if condition (12) is fulfilled from the definition of positive definite operator we have

$$\bar{F}[u] = \frac{1}{2} \langle N(u), KN(u) \rangle = \frac{1}{2} \langle v, Kv \rangle > 0 \tag{4.19}$$

for every $v \in D(K)$ and $v \neq \emptyset_V$. Since $\bar{F}[u_0] = 0$ it follows that $\bar{F}$ is minimum at u_0: property (d) is thus proved. (End of the proof.)

The linear operator

$$R(u; v) = N'^*_u(u; Kv) \tag{4.20}$$

transforms the given operator N into a potential operator $\bar{N}(u) = R(u; N(u))$ and is, therefore, an integrating operator.

The theorem then gives, under generous conditions on the operator N, an infinity of integrating operators, i.e. one for every operator K that fulfils the five conditions (7)–(11). This theorem then provides an extended variational formulation for nonlinear problems.

The relations between the various operators is shown in Fig. 3. Figure 4 shows the flow diagram of the operations to be performed in order to give a variational formulation to a given problem.

5. HOW TO CHOOSE THE INTEGRATING OPERATOR

The integrating operator for an operator N has the general form of (4.20). The integral operator K must be invertible, symmetric, wiht a range contained in $D(N'^*_u)$.

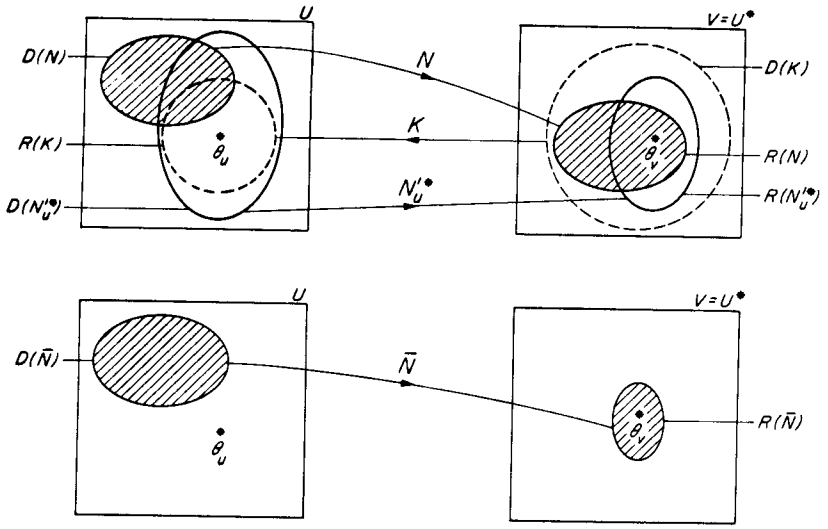


Fig. 3. Relation between the operators.

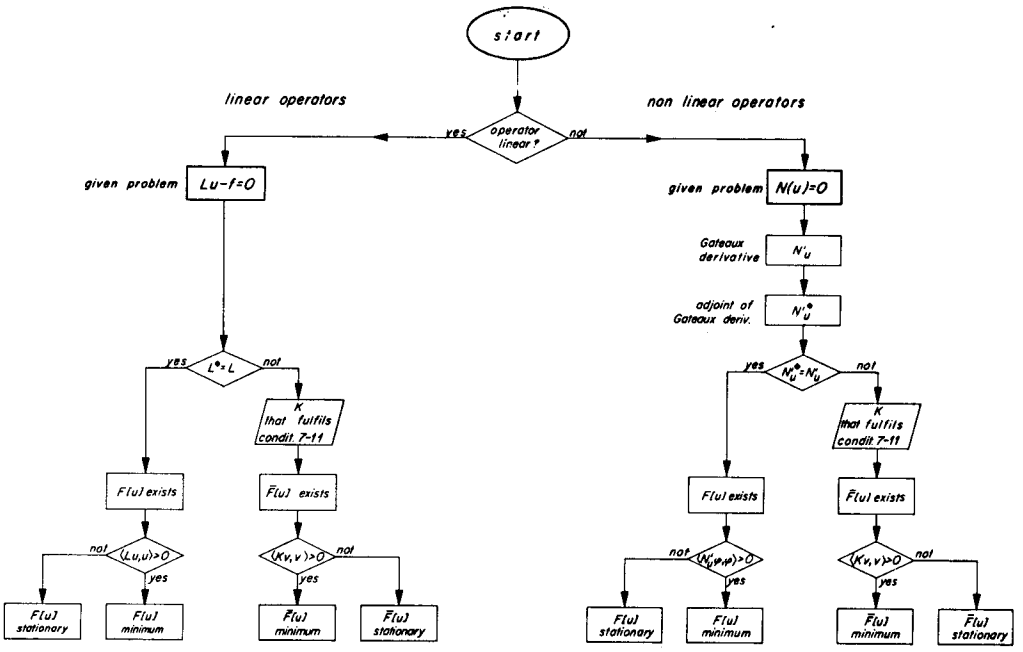


Fig. 4. The flow chart to give a variational formulation to a problem.

One source of such operators is the inverse of symmetric positive-definite differential operators: these are integral operators whose kernel is the Green function of the differential operator. So if one considers the operator

$$L = \left\{ -\frac{\mathrm{d}^2}{\mathrm{d}t^2}, u(0) = 0, u(T) = 0, u \in C^2(0, T) \right\} \tag{5.1}$$

which is symmetric positive-definite (and thus invertible), its inverse is

$$Kv = \int_0^T \left[-(t-\tau)H(t-\tau) + (T-\tau)\frac{t}{T} \right] v(\tau)\,\mathrm{d}\tau \tag{5.2}$$

where $H(z)$ is the unit-step Heaviside function.

In principle every Strum–Liouville formal differential operator

$$\mathscr{L} = -\frac{\mathrm{d}}{\mathrm{d}t}\left[p(t)\frac{\mathrm{d}}{\mathrm{d}t}\right] + q(t) \tag{5.3}$$

with $p(t) > 0$ and $q(t) \geqslant 0$ in $(0, T)$, associated with boundary conditions that make the operator L symmetric and invertible, gives rise to an operator $K = L^{-1}$ that is symmetric, invertible and positive definite. Table 1 gives some of these operators. The six Green functions of Table 1 are plotted in Fig. 5.

The use of Green function to form integral operators K for problems of several variables has two serious drawbacks.

The first is that Green functions for partial differential operators are known only for very simple domains, like circle, rectangle, half plane, etc. When one considers the great flexibility of the finite element methods for domains of arbitrary shape, this appears as a serious defect. The second is that even for simple domains the Green functions are expressed by series. In numerical applications one must truncate the series. Unfortunately the resulting kernel is degenerate and then K is not invertible. It is invertible only in the subspace spanned by the functions contained in the truncated kernel.

Fortunately we can easily find another method to determine an integral operator K. To show it we shall take an inductive approach.

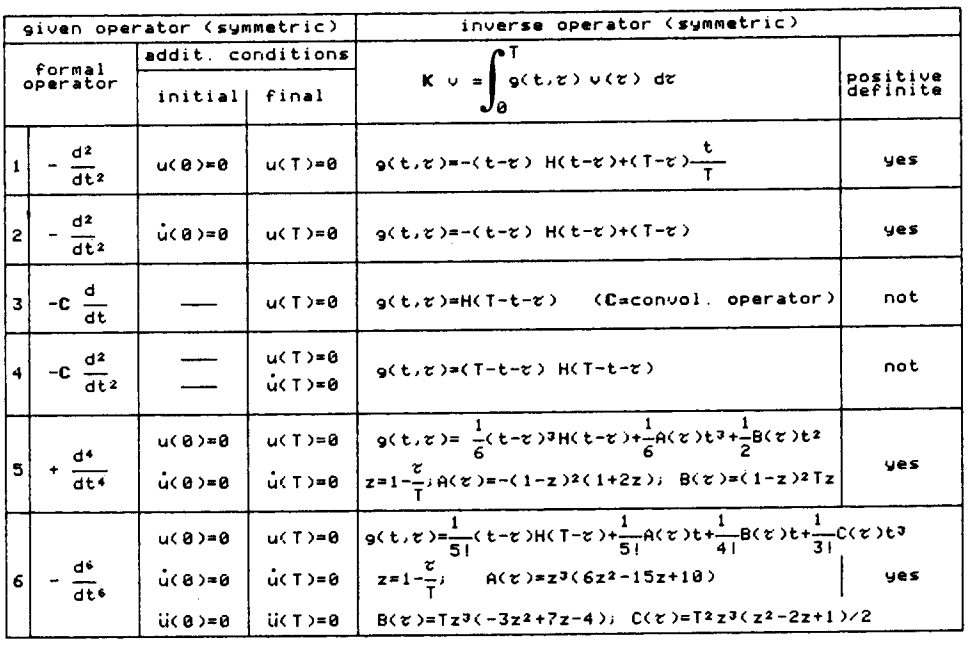
Let us consider at first problems in one variable, say t, and suppose that the interval is $(0, 1)$. Let $\varphi(t)$ be a function that belongs to the domain of $N_u'^*$. In practice, since $N_u'^*$ is a linear operator, the boundary/initial conditions associated with it must be homogeneous. Then $\varphi(t)$ must satisfy these homogeneous boundary/initial conditions. Let us suppose that these are $w(0) = 0$ and $w(1) = 0$. Let us consider the integral transform

$$w(t) = Kv = \int_0^1 f(t, \tau)\varphi(t)\varphi(\tau)v(\tau)\,\mathrm{d}\tau \tag{5.4}$$

where (up to now) $f(t, \tau)$ is an arbitrary function symmetric in the two variables t and τ. The function $w(t)$ satisfies the same homogeneous boundary conditions of $\varphi(t)$ and then it belongs to $D(N_u'^*)$. If we can find a function $f(t, \tau)$ that makes the operator K positive definite it will be surely invertible.

Table 1. Integral operators K for ordinary differential equations

given operator (symmetric)				inverse operator (symmetric)	
	formal operator	addit. conditions		$Kv = \int_0^T g(t,\tau)\,v(\tau)\,\mathrm{d}\tau$	positive definite
		initial	final		
1	$-\frac{\mathrm{d}^2}{\mathrm{d}t^2}$	$u(0)=0$	$u(T)=0$	$g(t,\tau)=-(t-\tau)\,H(t-\tau)+(T-\tau)\frac{t}{T}$	yes
2	$-\frac{\mathrm{d}^2}{\mathrm{d}t^2}$	$\dot{u}(0)=0$	$u(T)=0$	$g(t,\tau)=-(t-\tau)\,H(t-\tau)+(T-\tau)$	yes
3	$-C\,\frac{\mathrm{d}}{\mathrm{d}t}$	—	$u(T)=0$	$g(t,\tau)=H(T-t-\tau)$ (C=convol. operator)	not
4	$-C\,\frac{\mathrm{d}^2}{\mathrm{d}t^2}$	— —	$u(T)=0$ $\dot{u}(T)=0$	$g(t,\tau)=(T-t-\tau)\,H(T-t-\tau)$	not
5	$+\frac{\mathrm{d}^4}{\mathrm{d}t^4}$	$u(0)=0$ $\dot{u}(0)=0$	$u(T)=0$ $\dot{u}(T)=0$	$g(t,\tau)=\frac{1}{6}(t-\tau)^3H(t-\tau)+\frac{1}{6}A(\tau)t^3+\frac{1}{2}B(\tau)t^2$ $z=1-\frac{\tau}{T}$; $A(\tau)=-(1-z)^2(1+2z)$; $B(\tau)=(1-z)^2Tz$	yes
6	$-\frac{\mathrm{d}^6}{\mathrm{d}t^6}$	$u(0)=0$ $\dot{u}(0)=0$ $\ddot{u}(0)=0$	$u(T)=0$ $\dot{u}(T)=0$ $\ddot{u}(T)=0$	$g(t,\tau)=\frac{1}{5!}(t-\tau)H(T-\tau)+\frac{1}{5!}A(\tau)t+\frac{1}{4!}B(\tau)t+\frac{1}{3!}C(\tau)t^3$ $z=1-\frac{\tau}{T}$; $A(\tau)=z^3(6z^2-15z+10)$ $B(\tau)=Tz^3(-3z^2+7z-4)$; $C(\tau)=T^2z^3(z^2-2z+1)/2$	yes



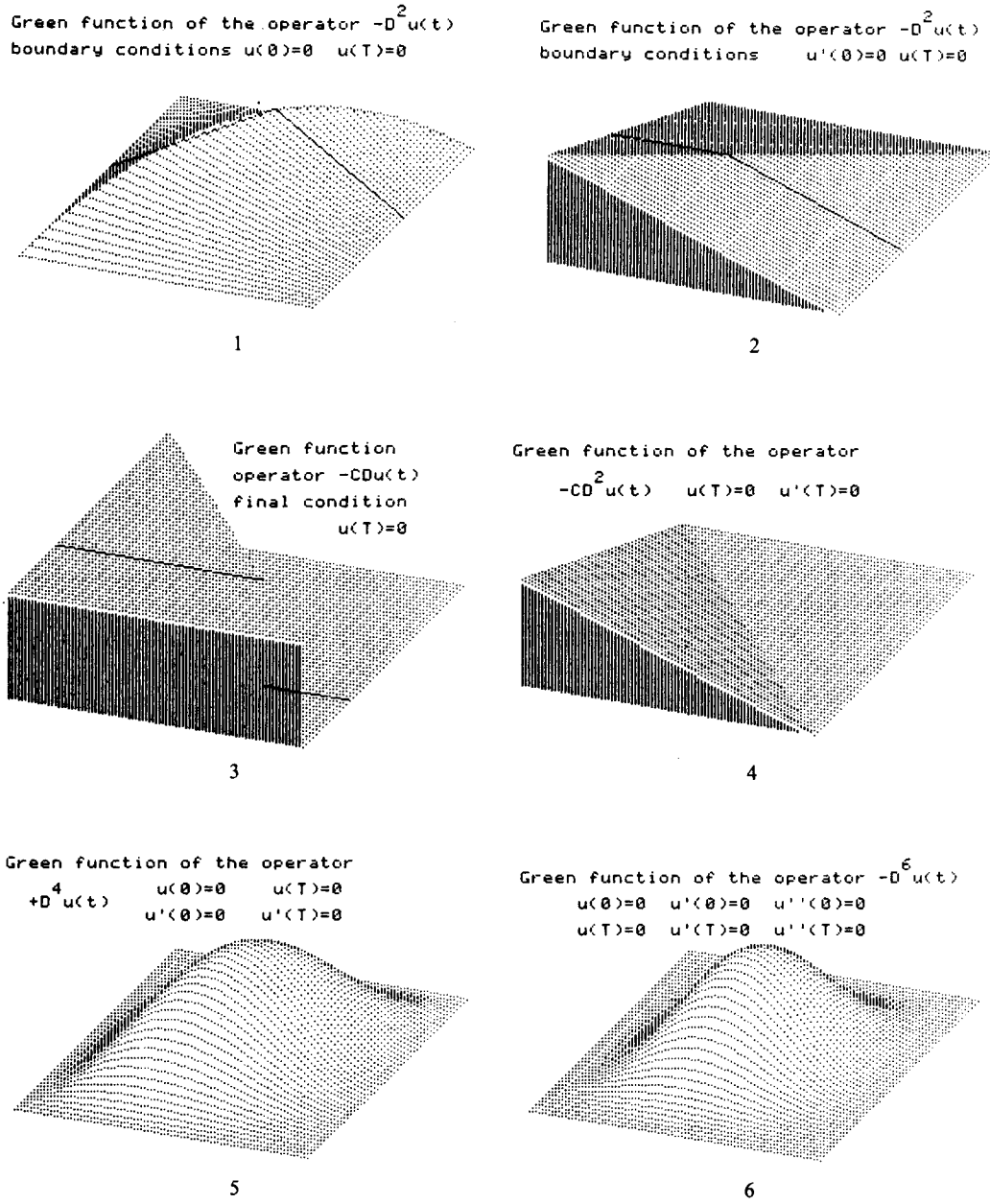


Fig. 5. Green functions to form the operation $\mathscr{K}$.

A first condition is that $f(t, \tau)$ must not be a polinomial; in this case the kernel would be degenerate. Let us take, e.g., the function $f(t, \tau) = \exp(t\tau)$. It will be

$$\begin{aligned}\langle v, Kv \rangle &= \int_0^1 v(t) \int_0^1 \exp(t\tau)\varphi(t)\varphi(\tau)v(\tau)\,\mathrm{d}\tau\,\mathrm{d}t \\ &= \sum_0^\infty \frac{1}{K!} \int_0^1 t^K v(t)\varphi(t)\,\mathrm{d}t \int_0^1 \tau^K v(\tau)\varphi(\tau)\,\mathrm{d}\tau \\ &= \sum_0^\infty \frac{1}{K!} (m_K)^2 \geqslant 0. \end{aligned} \tag{5.5}$$

Let us observe that the momenta m_K are simultaneously zero if and only if $v(t)$ is identically zero. It follows that for $v(t) \neq \emptyset_V$ is

$$\langle v, Kv \rangle > 0. \tag{5.6}$$

Thus we have found an integral operator that satisfies all conditions of Theorem 2.

In the case that the boundary/initial conditions involve the vanishing of the first derivatives, say $w'(0) = 0$ it is enough to choose a function $\varphi(t)$ that satisfies both the condition $\varphi'(0) = 0$ and $\varphi(0) = 0$. Since

$$w'(t) = \int_0^1 [f'\varphi(t) + f\varphi'(t)]\varphi(\tau)v(\tau)\,d\tau \tag{5.7}$$

it follows that $w'(0) = 0$.

For 2-dimensional problems an integral operator may be the following

$$Kv = \int_0^1 \int_0^1 \exp(x\xi y\eta)\varphi(x, y)\varphi(\xi, \eta)v(\xi, \eta)\,d\xi\,d\eta \tag{5.8}$$

where $\varphi(x, y)$ satisfies the homogeneous boundary/initial conditions of $D(N_u'^*)$.

This method of finding operators K is particularly suited for numerical applications.

Remark 2

If the operator N is linear and symmetric, so that a restricted variational formulation exists, the extended variational formulation contains it as a particular case. In fact if we choose $K = L^{-1}$ the functional $\bar{F}[u]$ is reduced to

$$\bar{F}[u] = \frac{1}{2}\langle Lu, KLu \rangle = \frac{1}{2}\langle Lu, u \rangle = F[u]. \tag{5.9}$$

Then for linear operators the extended variational formulation includes the restricted one as a particular case.

What happens if the operator K is potential but nonlinear? We need a generalization of Theorem 2 to include operators K that depend on u and are symmetric, i.e. such that

$$\langle \psi, K(u; \varphi) \rangle = \langle \varphi, K(u; \psi) \rangle. \tag{5.10}$$

Such a generalization is given in Appendix 2.

Remark 3

The condition that the range of K be contained in the domain of $N_u'^*$ si required to give meaning to the expression $N_u'^* KN(u)$. But if we start with the functional

$$\bar{F}[u] = \frac{1}{2}\langle N(u), KN(u) \rangle \tag{5.11}$$

and choose any K that is positive definite and is applicable to $N(u)$, we have a generalization of the least squares method. For $V = U$ and $K = I$ (I is the identity operator), we obtain the least squares method.

In conclusion, if we *give up* the condition $D(N_u'^*) \supset R(K)$, the functional of the extended variational formulation includes that of the least squares method as a particular case. We stress the fact that the extended variational formulation *is not* in itself a generalization of the least squares method because it requires the condition $D(N_u'^*) \supset R(K)$.

6. APPLICATIONS

To explain the procedure to be followed in practice, we here give two examples.

The procedure is divided into two steps. First, we test whether the given operator satisfies the six conditions of Theorem 2 and second, we select an operator K that satisfies the five conditions of Theorem 2 (and, if possible, the sixth one).

Example 1

A first order linear ordinary differential equation with initial condition.

Let us start with a very simple example. We want to give a variational formulation in the extended sense to the problem

$$\begin{cases} \mathrm{d}/\mathrm{d}t u(t) = f(t) & 0 \leqslant t \leqslant T \\ u(0) = 0 \qquad u \in C^1(0, T) & f \in C(0, T) \end{cases} \tag{6.1}$$

that does not admit of a classical variational formulation. It is an initial-value problem. As a first step, we must check whether the given operator satisfies the six conditions of Theorem 2.

Let us take as U the space $C^1(0, T)$ of continuously differentiable functions that vanish in the origin, and as V the space $C(0, T)$. The bilinear functional

$$\int_0^T u(t)v(t)\,\mathrm{d}t \tag{6.2}$$

puts the two spaces into duality. The norms

$$\|u\| = \sqrt{\int_0^T u^2(t)\mathrm{d}t} \qquad \|v\| = \sqrt{\int_0^T v^2(t)\mathrm{d}t} \tag{6.3}$$

induces topologies that make the bilinear functional continuous in both arguments.

If u_1 and u_2 are two functions of the domain, then also the convex combination

$$\eta(t) = \lambda u_1(t) + (1 - \lambda)u_2(t) \tag{6.4}$$

belongs to the domain. The domain is convex and thus, *a fortiori*, simply connected. In practice, every time we have linear (initial or boundary or mixed conditions) whether homogeneous or non-homogeneous, the domain is convex. Thus we may omit to check for the condition of simple connectivity in the following examples.

Let us perform the Gateaux derivative

$$N(u) = \left\{\frac{\mathrm{d}}{\mathrm{d}t}u,\ u(0) = 0,\ u \in C^1(0, T)\right\} \tag{6.5}$$

$$N'_u\varphi = \left\{\frac{\mathrm{d}}{\mathrm{d}t}\varphi,\ \varphi(0) = 0,\ \varphi \in C^1(0, T)\right\} \tag{6.6}$$

$$N'^*_u\psi = \left\{-\frac{\mathrm{d}}{\mathrm{d}t}\psi,\ \psi(T) = 0,\ \psi \in AC(0, T)\right\}. \tag{6.7}$$

The Gateaux derivative thus exists. Its domain is formed of the functions of class $C^1(0, T)$: they form a dense subset of U with respect to the topology induced by norm (6.3) (in this case the domain coincides with U).

Since the adjoint homogeneous problem

$$-\frac{\mathrm{d}}{\mathrm{d}t}\psi = 0 \quad \psi(T) = 0 \tag{6.8}$$

has only the null solution, the operator $N_u'^*$ is invertible. We thus see that the first six conditions of Theorem 2 are fulfilled.

If we choose an operator K whose kernel is one of the Green functions of Table 1, we satisfy conditions 7–11 of Theorem 2. In fact, the domain of K is that of measurable functions and thus contains $R(N)$. On account of the final conditions $g(T, t) = 0$, the transformed function Kv satisfies the final condition $\psi(T) = 0$. The operator K is obviously linear, invertible and symmetric.

In particular, the Green functions 1, 2, 5, 6 make the operator K positive definite and thus satisfy condition 12 of Theorem 2.

If we consider the Green functions of row 2, which make the operator K positive-definitive, we obtain the integrating operator

$$Rv = -\frac{\mathrm{d}}{\mathrm{d}t}\int_0^T g(t,\tau)v(\tau)\,\mathrm{d}\tau = \int_0^T -\frac{\partial g(t,\tau)}{\partial t}\,v(\tau)\mathrm{d}\tau$$

$$= \int_0^T H(t-\tau)v(\tau)\,\mathrm{d}\tau = \int_0^t v(\tau)\,\mathrm{d}\tau \tag{6.9}$$

which is simply the indefinite integral! If we apply it to Problem (6.1) we obtain the problem

$$u(t) - \int_0^t f(\tau)\,\mathrm{d}\tau = 0 \tag{6.10}$$

that has the same solution as the given one. In this very simple example, we have obtained the solution directly.

To find the functional of the linear problem (6.8), we may proceed directly by multiplying by δu and integrating. We, therefore, obtain

$$\bar{F}[u] = \int_0^T \left[\frac{1}{2}u^2(t) - u(t)\int_0^t f(\tau)\,\mathrm{d}\tau\right]\mathrm{d}t. \tag{6.11}$$

This functional may also be obtained by the general procedure explained in Section 6, choosing $\eta(\lambda) = \lambda u$. The same functional may be obtained by starting with the general form (4.11), i.e.

$$\bar{F}[u] = \frac{1}{2}\int_0^T \left[\frac{\mathrm{d}}{\mathrm{d}t}u(t) - f(t)\right]\int_0^T g(t,\tau)\left[\frac{\mathrm{d}}{\mathrm{d}\tau}u(\tau) - f(\tau)\right]\mathrm{d}\tau\,\mathrm{d}t \tag{6.12}$$

and performing the integration by parts, assuming the initial condition $u(0) = 0$.

In conclusion the linear initial value problem (6.1) admits of an extended variational formulation: its solution is the minimum of the functional (6.11).

We may remark that the integrating operator R has the form stated in eqn (1.7).

We also observe that if one uses the Green function of the convolution operator (row 3 in Table 1), the operator K is not positive and the corresponding functional is not minimum at the solution.

Example 2

A nonlinear ordinary differential equation of the first order with initial condition.

Let us consider the problem

$$\dot{u}(t) = f(t; u(t)) \quad u(0) = a \quad u \in C^1(0, T) \tag{6.13}$$

where f is an assigned function. Since this is a Cauchy problem, the existence and the

uniqueness of the solution is assured under the usual hypothesis. We have

$$N(u) = \left\{ \frac{\mathrm{d}}{\mathrm{d}t} u(t) - f(t; u(t)),\ u(0) = a,\ u \in C^1(0, T) \right\} \tag{6.14}$$

$$N'_u \varphi = \left\{ \frac{\mathrm{d}}{\mathrm{d}t} \varphi(t) - \frac{\partial f}{\partial u} \varphi(t),\ \varphi(0) = 0,\ \varphi \in C^1(0, T) \right\} \tag{6.15}$$

$$N'^*_u \psi = \left\{ -\frac{\mathrm{d}}{\mathrm{d}t} \psi(t) - \frac{\partial f}{\partial t} \psi(t),\ \psi(T) = 0,\ \psi \in AC(0, T) \right\}. \tag{6.16}$$

The adjoint homogeneous problem is

$$-\frac{\mathrm{d}}{\mathrm{d}t} \psi(t) - \frac{\partial f}{\partial u} \psi(t) = 0 \quad \psi(T) = 0. \tag{6.17}$$

The equation is linear with variable coefficients containing $u(t)$. It can be easily seen that it has only the null solution, and thus the operator N'^*_u is invertible. The conditions 1–6 of Theorem 2 are satisfied.

If we choose any Green function from Table 1, we have a variational formulation of Problem (6.13). If we let

$$h(t; u(t), \dot{u}(t)) = \dot{u}(t) - f(t; u(t)), \tag{6.18}$$

we obtain the functional

$$\bar{F}[u] = \frac{1}{2} \int_0^T h(t; u(t), \dot{u}(t)) \int_0^T g(t, \tau) h(t; u(\tau), \dot{u}(\tau)\, \mathrm{d}\tau\, \mathrm{d}t. \tag{6.19}$$

Choosing Green function 2 from Table 1, the integrating operator becomes

$$R(u, v) = \left[-\frac{\mathrm{d}}{\mathrm{d}t} - \frac{\partial f}{\partial u} \right] \int_0^T [-(t - \tau) H(t - \tau) + (T - \tau)] v(\tau)\, \mathrm{d}\tau, \tag{6.20}$$

which after an integration by parts becomes

$$R(u, v) = \int_0^T v(\tau)\, \mathrm{d}\tau - \frac{\partial f}{\partial u} \int_0^T [-(t - \tau) H(t - \tau) + (T - \tau)] v(\tau)\, \mathrm{d}\tau. \tag{6.21}$$

It is a linear operator on the argument $v(t)$ with coefficients that contain $u(t)$.

One of the advantages of the presence of the operator K is that in the functional one can remove the higher order derivatives by integrating by parts. The kernel $g(t, \tau)$ "absorbs" the derivatives. By contrast, with the least squares method, this elimination is not possible. An expression like

$$\int_0^T \dot{u}(t) \int_0^T g(t, \tau) \dot{u}(\tau)\, \mathrm{d}t \tag{6.22}$$

is equivalent to

$$\int_0^T u(t) \int_0^T \frac{\partial^2 g(t, \tau)}{\partial t \partial \tau} u(\tau)\, \mathrm{d}\tau\, \mathrm{d}t + A - 2B \tag{6.23}$$

$$A = [u(t)\, [g(t, \tau) u(\tau)]_{\tau = 0}^{\tau = T}]_{t = 0}^{t = T} \tag{6.24}$$

$$B = \int_0^T u(t)\left[\frac{\partial g(t,\tau)}{\partial t}u(\tau)\right]_{\tau=0}^{\tau=t} \mathrm{d}t \tag{6.25}$$

as can be seen by performing two integrations by parts.

With a proper choice of the Green function one may obtain $A = 0$ and $B = 0$ (for example, the Green functions 1, 5, 6 of Table 1). Such a choice the functional (6.19) becomes with

$$\bar{F}[u] = \frac{1}{2}\int_0^T u(t)\int_0^T \frac{\partial^2 g}{\partial\tau\partial t}u(\tau)\,\mathrm{d}\tau\,\mathrm{d}t + \int_0^T f(t;u(t))\int_0^T \frac{\partial g}{\partial\tau}u(\tau)\,\mathrm{d}t\,\mathrm{d}\tau + \frac{1}{2}\int_0^T f(t;u(t))\int_0^T g(t,\tau)f(\tau;u(\tau))\,\mathrm{d}\tau\,\mathrm{d}t. \tag{6.26}$$

In the following section we shall give a numerical solution to this problem.

7. NUMERICAL EXPERIMENTS

In order to check on the numerical performance of the variational formulation obtained, we have tested it on some equations of different kinds.

These experiments are only a taste of some possible applications: a systematic comparison with existing numerical methods is left to more specialized authors.

In every test, we used integral operators K that assure the minimum of the functional. Ritz' method of undetermined coefficients was used. To find the minimum of the functional, we used the second order gradient method, which approximates the functional with a "paraboloid" at every point ([1], p. 64).

The procedure is the following. We approximate the solution with the linear combination

$$\bar{u}(t) = \varphi_0(t) + \sum_K a_K\varphi_K(t) \tag{7.1}$$

where $\varphi_0(t)$ satisfies the given non-homogeneous conditions and the $\varphi_K(t)$ satisfies the corresponding homogeneous conditions. Letting $\mathbf{a} = (a_1, a_2, \ldots a_n)$ and denoting the residual by $r(u(t))$, every functional, with our choice of K, has the form of

$$\bar{F}[u] = \frac{1}{2}\int_0^T r(\bar{u}(t))\int_0^T g(t,\tau)r(\bar{u}(\tau))\,\mathrm{d}\tau\,\mathrm{d}t \tag{7.2}$$

and since u is a function of $\mathbf{a}$, $\bar{F}$ is reduced to the function

$$f(\mathbf{a}) = \frac{1}{2}\int_0^T t(t;\mathbf{a})\int_0^T g(t,\tau)r(\tau;\mathbf{a})\,\mathrm{d}\tau\,\mathrm{d}t \tag{7.3}$$

which is quadratic on the residual. When the equation is linear, the residual is linear in $\mathbf{a}$: it follows that the function $f(\mathbf{a})$ is quadratic on the coefficients.

By selecting an initial vector $\mathbf{a}$, we are going to find the vector that makes the function $f(\mathbf{a})$ minimum. The variation in $f(\mathbf{a})$ for a variation in the coefficients is given (in second order approximation) by

$$\Delta f(\mathbf{a}) = \sum_K g_K(\mathbf{a})\Delta a_K + \sum_{hK}\frac{1}{2}e_{hK}(\mathbf{a})\Delta a_h\Delta a_K \tag{7.4}$$

from which, considering the symmetry of $g(t,\tau)$, it follows that

$$g_K(\mathbf{a}) = \int_0^T \frac{\partial r(t,\mathbf{a})}{\partial a_K}\int_0^T g(t,\tau)r(\tau;\mathbf{a})\,\mathrm{d}\tau\,\mathrm{d}t \tag{7.5}$$

$$e_{hK}(\mathbf{a}) = \int_0^T \frac{\partial r(t, \mathbf{a})}{\partial a_K} \int_0^T g(t, \tau) \frac{\partial r(\tau_1, \mathbf{a})}{\partial a_K} \, d\tau \, dt. \tag{7.6}$$

The vector $\mathbf{g}$ and the matrix $\mathbf{e}$ are the gradient and the hessian of the function $\Delta f(\mathbf{a})$, respectively. The maximum decrease in the variation $\Delta f(\mathbf{a})$ arises for the variation of $\mathbf{a}$ that satisfies the system

$$\sum_K e_{hK}(\mathbf{a}) \Delta a_K + g_h(\mathbf{a}) = 0. \tag{7.7}$$

By solving the system, we find the Δa_K and with it the new vector

$$\bar{a}_K = a_K + \Delta a_K. \tag{7.8}$$

This procedure will be repeated until, say, the first four digits remain fixed.

We have tested 5 ordinary differential equations of the first and second orders, each one associated with initial conditions, both homogeneous and non-homogeneous. None of these problems admits of a variational formulation in the restricted sense (i.e. in the classical sense). These 5 problems are summarized in Table 2, and then solutions are summarized in Fig. 6.

On account of their preliminary value, the calculations have been made with a personal computer (an Apple II). The graphs are the hard copy of the graphic page.

Each picture reports both the exact solution and the graph of the error (with a different scale). We have used the Ritz method with 5 coefficients and powers of t as trial functions. The integrations have been performed with Gaussian quadrature formulas (10 points).

An interesting point must be noted: in solving second-order equations we have obtained good results even by using Green functions that do not satisfy the initial conditions. For example, in Problem 4 (Table 2), if one uses Green function 2 (Table 1), which does not satisfy the boundary condition $\dot{u}(0) = 0$, the result is even better than that obtained with the Green function 5 (Table 1). This is a consequence of the fact that when K is only a positive definite operator that does not satisfy the condition $D(N_u'^*) \supset R(K)$, the functional is a generalization of the least squares method.

We have also performed some tests by means of the finite element method, obtaining good results. The main novelty is a complication arising from the integro-differential nature of the modified equation. When one evaluates the contribution of every element to the whole integral, every element shows the contribution of all the other elements, and not only of the adjacent ones. It follows that the stiffness matrix is no longer banded, as it usually is.

CONCLUDING REMARKS

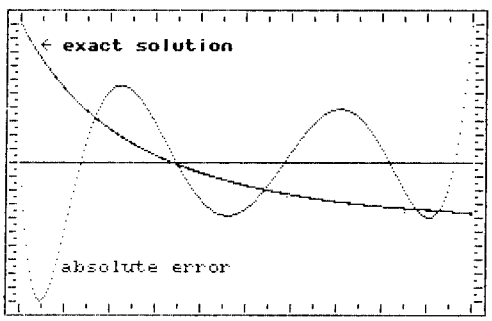
One may say that the integro-differential nature of the equation is the price we have to pay to give a variational formulation to equations that do not admit of one in the restricted sense.

Theorem 2 gives both the general form of the integrating operator and the explicit form of the functional.

Table 2. Selected problems chosen to test the method

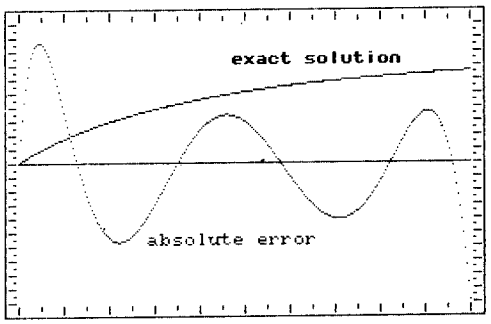
	differential equation	condition	exact solution
1	$u' = 0.1u - 2\frac{\exp(-0.1t)}{(1+t)^2}$	$u(0)=1$	$u = \exp(-0.1t)\frac{1-t}{1+t}$
2	$u' = \exp(-t)/(1+u)$	$u(0)=0$	$u = -1 + \sqrt{3-2\exp(-t)}$
3	$u' = \frac{2t}{1+\exp(t)} - \frac{\exp(t)}{(1+t)^2}u^2$	$u(0)=0.5$	$u = \frac{1+t^2}{1+\exp(t)}$
4	$u'' + u = 0$	$u(0)=0$ $u'(0)=1$	$u = \sin(t)$
5	$u''-u'-(3-t)\exp(-2t)u^3=0$	$u(0)=1$ $u'(0)=-1$	$u = \exp(t)/(1-t)$

```
PROGRAM NAME:
620  REM        VAR 5 III

EXACT SOLUTION:
430 UE(K) =  EXP ( - 0.1 * T) * (
     1 - T) / (1 + T)

RESIDUAL:
170 R = DU + 0.1 * U + 2 *  EXP (
      - 0.1 * T) / ((1 + T) * (1 +
     T))

INTERVAL:  0,3           INITIAL VALUE:     1
GREEN FUNCTION:      2   GAUSS POINTS:      10
COEFFICIENTS:        5   COEFF. TOLERANCE: 1E-04
FUNCTIONAL MINIMUM: 5.24887685E-06
MAX ABSOLUTE ERROR: 6.31047506E-03
A(1)= -1.99041328        A(2)= 1.57591146
A(3)= -.763893787        A(4)= .198751639
A(5)= -.020832488
```



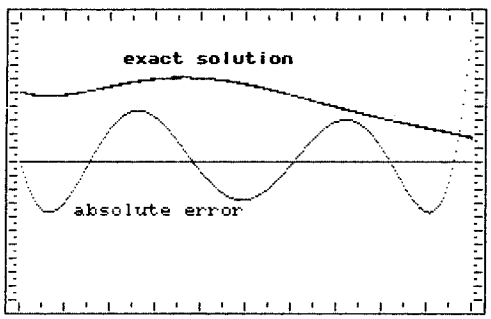
a

```
PROGRAM NAME:
620  REM      VAR 5 I

EXACT SOLUTION:
430 UE(K) =  - 1 +  SQR (3 - 2 *
      EXP ( - T))

RESIDUAL:
170 R = DU -  EXP ( - T) / (U + 1


INTERVAL:  0,2           INITIAL VALUE:     0
GREEN FUNCTION:      2   GAUSS POINTS:      10
COEFFICIENTS:        5   COEFF. TOLERANCE: 1E-04
FUNCTIONAL MINIMUM: 1.56546449E-07
MAX ABSOLUTE ERROR: 9.47717112E-04
A(1)= .978100362         A(2)= -.805386958
A(3)= .473815226         A(4)= -.167117925
A(5)= .0250540013
```



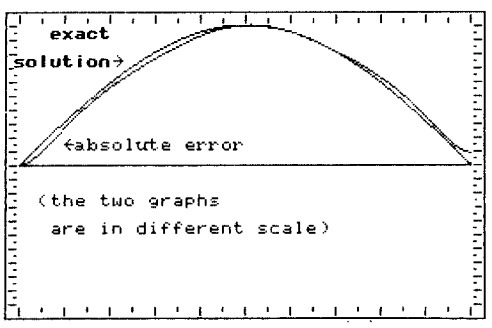
b

```
PROGRAM NAME:
620  REM        VAR 5 II

EXACT SOLUTION:
430 UE(K) = (1 + T * T) / (1 +  EXP
     (T))

RESIDUAL:
170 R = DU - 2 * T / (1 +  EXP (T
     )) +  EXP (T) / (1 + T * T) *
     U * U

INTERVAL:  0,5           INITIAL VALUE:     .5
GREEN FUNCTION:      2   GAUSS POINTS:      10
COEFFICIENTS:        5   COEFF. TOLERANCE: 1E-06
FUNCTIONAL MINIMUM: 7.92102174E-06
MAX ABSOLUTE ERROR: .0132557305
A(1)= -.218696872        A(2)= .450533836
A(3)= -.242312676        A(4)= .0487751475
A(5)= -3.42543548E-03
```



c

```
PROGRAM NAME:
620  REM   VAR SEC II (G=5)

EXACT SOLUTION:
430 UE(K) =  SIN (T)

RESIDUAL:
170 R = D2 + U

INTERVAL:  0,3.14159     INITIAL VALUE:     0
GREEN FUNCTION:      5   GAUSS POINTS:      10
COEFFICIENTS:        5   COEFF. TOLERANCE: 1E-04
FUNCTIONAL MINIMUM: 1.22808961E-11
MAX ABSOLUTE ERROR: 3.90659552E-04
A(1)= -2.14762539E-03    A(2)= -.161653378
A(3)= -5.80557105E-03    A(4)= .0120451442
A(5)= -1.27804284E-03
```

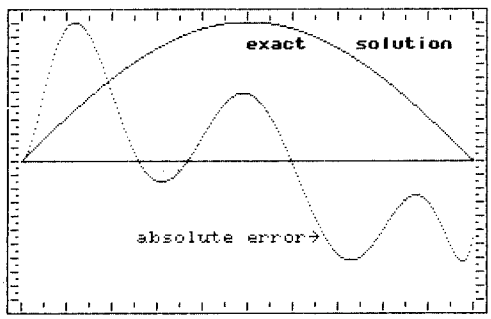


d

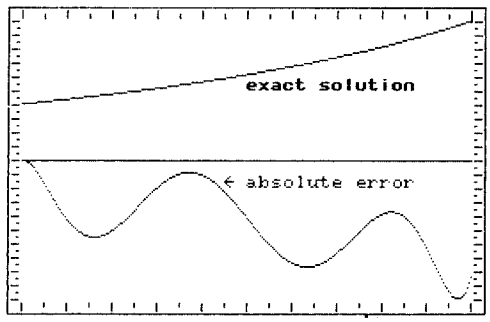
Fig. 6.

```
PROGRAM NAME:
620  REM     VAR SEC II

EXACT SOLUTION:
430 UE(K) =  SIN (T)

RESIDUAL:
170 R = D2 + U

INTERVAL:  0,3.14159     INITIAL VALUE:     0
GREEN FUNCTION:      2   GAUSS POINTS:      10
COEFFICIENTS:        5   COEFF. TOLERANCE: 1E-04
FUNCTIONAL MINIMUM: 2.11363562E-09
MAX ABSOLUTE ERROR: 2.67057913E-05
A(1)= -9.76962197E-04    A(2)= -.163122323
A(3)= -4.9998196E-03     A(4)= .0118300141
A(5)= -1.25579596E-03
```

```
PROGRAM NAME:
620  REM       VAR SEC I

EXACT SOLUTION:
430 UE(K) =  EXP (T) / (1 - T)

RESIDUAL:
170 R = D2 - D1 - (3 - T) *  EXP
     ( - 2 * T) * U * U * U

INTERVAL:  0,.4          INITIAL VALUE:     1
GREEN FUNCTION:      2   GAUSS POINTS:      10
COEFFICIENTS:        5   COEFF. TOLERANCE: 1E-04
FUNCTIONAL MINIMUM: 2.28666592E-07
MAX ABSOLUTE ERROR: 2.24309042E-05
A(1)= 2.5150128          A(2)= 2.35933726
A(3)= 5.0587168          A(4)= -5.38084805
A(5)= 14.3043253
```



e



f

Fig. 6. Apple II.

The method seems particularly promising for the equations of fluid-dynamics, because it is applicable both to evolution equations and to nonlinear equations.

It must be stressed that this method provides a general procedure to find the functional whatever the boundary conditions might be, whether homogeneous or not.

One may be disappointed by the fact that the form of the equation is greatly altered by the application of an integrating operator. Without any doubt, the change in form obscures the search for peculiar properties of the solution. Traditional criteria used to identify properties of the solution from the structure of the equation become useless. But *from a physical point of view, what is essential is not the form of the equation but the solution.* We describe physical laws in differential form because this is today the most simple and direct way. But the transformation we have performed on the equation changes the form, not the content, of the equation. The solution is not changed, and this is the important fact.

With Theorem 2 we have shown that the variational characterization depends on the form of the equation: it is always possible to change the form of the equation in order to obtain a variational formulation. Thus, the characterization of a physical phenomenon according to its admitting of a variational formulation or not is a formal criterion, not a substantial one.

To stress this point, let us compare two classical variational formulations: the Dirichlet variational principle for the Poisson equation and the Hamilton principle for the equation of the motion of a particle. In the first case the boundary conditions are an essential part of the variational formulation: the solution of the Poisson equation makes the functional minimum. The solution may be found by applying the direct methods of the Calculus of Variations.

In the second case, the initial conditions are not an essential part of the variational formulation: one has to ignore the physical condition of the initial velocity and must add a fictitious final condition. The natural motion is privileged by the stationary property of the action functional in a class of functions that satisfy one given initial condition and one fictitious final condition. The physics of the problem is violated. A typical initial value

problem is artificially changed into a boundary value problem in order to obtain a variational formulation. The result is that one cannot use the Hamilton principle to obtain the solution by direct methods. The added final condition is not known.

This comparison shows that the characterization of physical phenomena based on the existence or not of a variational formulation, even in a classical context, masks substantial differences. Classically, the variational formulation is a property shared by some boundary value problems. Initial value problems cannot be encompassed in it. One must artificially change an initial value problem into a boundary value one in order to obtain the variational formulation.

If we insist on giving a variational formulation to initial value problems, it is more natural to change the form of the equation, preserving both the solution and the given physical initial conditions. In this way one may find functionals that attain their minimum at the solution, and the solution can be found by direct methods.

Acknowledgement—The author is grateful to the engineering student F. Beltram for his invaluable help in the editing of the paper.

REFERENCES

[1] P. R. ADBY and M. A. H. DEMPSTER, *Introduction to Optimization Methods*. Chapman and Hall, London (1974).
[2] I. M. ANDERSON and T. DUNCHAMP, *J. Math.* **102**, 781–868 (1980).
[3] R. W. ATHERTON and G. M. HOMSY, *Stud. Appl. Math.* **54**, 31–60 (1975).
[4] L. COLLATZ, *The Numerical Treatment of Differential Equations*. Springer, Berlin (1960).
[5] D. R. DAVIS, *Trans. Am. Math. Soc.* **30**, 716, 736 (1928).
[6] P. DEDECKER, *Bull. Acad. Roy. Belg. Sc.* **36**, 63–70 (1950).
[7] P. DEDECKER and W. M. TULCZYJEW, *Lect. Notes Math.* Vol. 836 (1979).
[8] V. P. DIDENKO, *Dokl. Akad. Nauck. SSSR* **240**, 736–740 (1978).
[9] J. DOUGLAS, *Trans. Am. Math. Soc.* **50**, 71–128 (1941).
[10] G. C. EVANS, *Am. Math. Soc.* (1918) (Reprinted by Dover, New York in 1964).
[11] B. A. FINLAYSON, *Phys. Fluids* **15**, 963–967 (1972).
[12] B. A. FINLAYSON and L. E. SCRIVEN, *J. Heath Mass Transfer* **10**, 799–821 (1957).
[13] Y. C. FUNG, *Foundation of Solid Mechanics*. Prentice-Hall, New York (1965).
[14] M. E. GURTIN, *Q. Appl. Math.* **22**, 252–256 (1964).
[15] M. E. GURTIN, *Arch. Rat. Mech. Anal.* **13**, 179–197 (1963).
[16] P. HAVAS, *Suppl. Nuovo Cim.* **5**, 363–388 (1957).
[17] H. HERRERA and J. BIELAK, *Arch. Rat. Mech. Anal.* **53**, 131, 149 (1974).
[18] H. HELMHOLTZ, Ueber die physikalische bedeutung des Prinzips der kleinsten Wirkung, *J. fuer die reine und angew. Matem.* 137–166, 213–222 (W.A.III, pp. 203–248) (1886).
[19] A. HIRSCH, *Math. Ann.* **49**, 49–72 (1897).
[20] I. HLAVACEK, *Aplik. Matem.* **4**, 278–297 (1969).
[21] M. KERNER, *Ann. of Math.* **34**, 546–572 (1933).
[22] C. LANCZOS, *Linear Differential Operators*. Van Nostrand, Amsterdam (1961).
[23] F. MAGRI, *Ann. Phys.* **99**, 334–344 (1976).
[24] F. MAGRI, *Int. J. Engng Sci.* **12**, 537–549 (1974).
[25] S. G. MICHLIN, *Variational Methods in Mathematical Physics*. Pergamon Press, Oxford (1964).
[26] S. G. MICHLIN, *The Problem of the Minimum of a Quadratic Functional*. Holden Day, New York (1965).
[27] M. MORSE, *The Calculus of Variations in the Large*, pp. 173–183. Verhandlungen des Internationalen Mathematiker-kongress, Zurich (1932).
[28] M. MORSE, *Bull. Am. Math. Soc.* **35**, 38–54 (1929).
[29] M. MORSE and H. FESHBACH, *Methods of Theoretical Physics*. McGraw-Hill, New York (1953).
[30] M. Z. NASHED, *Nonlinear Functional Analysis and Applications* (Edited by Rall). Academic Press, New York.
[31] J. L. NOWINSKI, *Int. J. Engng Sci.* **19**, 1377–1390 (1981).
[32] J. T. ODEN and J. N. REDDY, *Variational Methods in Theoretical Mechanics*. Springer Verlag, Berlin (1976).
[33] P. J. OLVER, *Math. Proc. Camb. Phil. Soc.* **88**, 71–88 (1980).
[34] R. REISS and E. J. HAUG, *Int. J. Engng Sci.* **16**, 231–251 (1978).
[35] V. M. SALOV, *Sov. Math. Dokl.* **iv**, 1046–1048 (1963).
[36] R. M. SANTILLI, *Foundations of Theoretical Mechanics I: The Inverse Problem of Newtonian Mechanics*. Springer Verlag, Berlin (1978).
[37] H. SCHAEFER, *Topological Vector Spaces*. Springer Verlag, Berlin (1978).
[38] F. TAKENS, *J. Diff. Geom.* **14**, 543–562 (1979).
[39] J. J. TELEGA, *J. Inst. Maths. Applic.* **24**, 175–195 (1979).
[40] T. TONTI, *Hadronic J.* **5**, 1404–1450 (1982).
[41] E. TONTI, *Rend. Acc. Lincei* **LII**, 175–181, 350–356 (1972).
[42] E. TONTI, *Rend. Acc. Lincei* **VL**, 293–300 (1968).

[43] E. TONTI, *Rend. Acc. Lincei* **LII**, 39–56 (1972).
[44] E. TONTI, *Seminario Matematico e Fisico di Milano* **XLVI** (1976), 163–257 (1976) (preprint: On the Formal Structure of Physical Theories, 1975, Consiglio Nazionale dell ricerche, 1975).
[45] E. TONTI, *Appl. Math. Modell.* **i**, 37–60 (1976).
[46] E. TONTI, *Annali di Matem. Pura Appl.* **XCV**, 331–360 (1972).
[47] E. TONTI, *Bull. Acad. Roy. de Belgique*, Scr. 5 **LV**, 137–165 (first part), 262–278 (second part) (1969).
[48] W. M. TULCZYJEW, *Bull. Soc. Math. France* **105**, 419–431 (1977).
[49] W. M. TULCZYJEW, *Bull. Acad. Pol. Sc. Ser. Math. Astron. Phys.* **24**, 1089–1096 (1976).
[50] M. M. VAINBERG, *Variational Methods for the Study of Nonlinear Operators*. Holden–Day, New York (1964).
[51] A. L. VANDERBAUWHEDE, Potential Operators and the Inverse Problem of Classical Mechanics, *Hadronic J.* **1**, 1177–1197 (1978).
[52] V. VOLTERRA, *Leçons sur les fonctions de ligne*. Gauthier Villars, Paris (1913).
[53] V. VOLTERRA, *Rend. Acc. Lincei* **III**, 97–105 (I), 141–148 (II), 153–158 (III) (1887).
[54] V. VOLTERRA, *Theory of Functionals and Integrodifferential Equations* London (1929) (reprinted in 1959 by Dover, New York).

(*Received* 3 *September* 1983)

APPENDIX 1

We propose to show that the operator $\tilde{N}$ defined by the relation

$$\tilde{N}(u) = N'^{*}_{u}(u; KN(u)) \tag{A1.1}$$

has a symmetric Gateaux derivative. To this end, we may consider it to be a composed operator

$$\tilde{N}(u) = N'^{*}_{u}(u; \psi) \quad \text{with} \quad \psi = KN(u). \tag{A1.2}$$

The total Gateaux differential is

$$\tilde{N}'_{u}(u; \delta u) = (N'^{*}_{u})'_{u}(u; \psi, \delta u) + (N'^{*}_{u})'_{\psi}(u; \delta\psi). \tag{A1.3}$$

To perform the adjoint, we make the scalar product

$$\langle \tilde{N}'_{u}(u; \delta u), \varphi \rangle = \langle (N'^{*}_{u})'_{u}(u; \psi, \delta u), \varphi \rangle + \langle (N'^{*}_{u})'_{\psi}(u; \delta\psi), \varphi \rangle. \tag{A1.4}$$

From the relation defining the adjoint

$$\langle N'_{u}(u; \varphi), \psi \rangle = \langle N'^{*}_{u}(u; \psi), \varphi \rangle \tag{A1.5}$$

one obtains, by partial differentiation in u, the relation

$$\langle N''_{uu}(u; \varphi, \delta u), \psi \rangle = \langle (N'^{*}_{u})'_{u}(u; \psi, \delta u), \varphi \rangle \tag{A1.6}$$

and in ψ

$$\langle N'_{u}(u; \varphi), \delta\psi \rangle = \langle (N'^{*}_{u})'_{\psi}(u; \delta\psi), \varphi \rangle, \tag{A1.7}$$

With these identities Relation (1.4) becomes

$$\langle \tilde{N}'_{u}(u; \delta u), \varphi \rangle = \langle N''_{uu}(u; \varphi, \delta u), \psi \rangle + \langle N'_{u}(u; \varphi), \delta\psi \rangle. \tag{A1.8}$$

Since

$$\delta\psi = KN'_{u}(u; \delta u) \tag{A1.9}$$

the last scalar product becomes

$$\langle N'_{u}(u; \varphi), KN'_{u}(u; \delta u) \rangle. \tag{A1.10}$$

Then

$$\langle \tilde{N}'_{u}(u; \delta u), \varphi \rangle = \langle N''_{uu}(u; \varphi, \delta u), \psi \rangle + \langle N'_{u}(u; \varphi), KN'_{u}(u; \delta u) \rangle. \tag{A1.11}$$

Let us observe that the second derivative of a non linear operator is symetric

$$N''_{uu}(u; \varphi, \psi) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\frac{\mathrm{d}}{\mathrm{d}\nu} N(u + \epsilon\varphi + \nu\psi)\Big|_{\substack{\epsilon=0 \\ \nu=0}} = N''_{uu}(u; \psi, \varphi). \tag{A1.12}$$

Besides, if we remember the symmetry of K we can say that the two scalar products of the second member are

symmetric in φ and δu. It follows that

$$\langle \bar{N}'_u(u; \delta u), \varphi \rangle = \langle \bar{N}'_u(u; \varphi), \delta u \rangle. \tag{A1.13}$$

Thus the Gateaux derivative on $\bar{N}$ is symmetric.

APPENDIX 2

Theorem 2 may be even more generalized by assuming that the operator K, while remaining linear and symmetric, contains u. It will be denoted as by $K(u; .)$. Such a generalization has the purpose of making the variational formulation of a nonlinear potential operator into a particular case of the extended formulation.

Theorem 3

If the operator $K(u; .)$ *satisfies the same conditions as Theorem* 2 *and the additional condition*

$$K'_u(u; N'_u(u; \varphi), \psi) = K'_u(u; N'_u(u; \psi), \varphi) \tag{A2.1}$$

then the operator

$$\bar{N}(u) = N'^*_u(u; K(u; N(u))) = N'^*_u K N(u) \tag{A2.2}$$

is potential.

Proof. Let us consider $\bar{N}$ as a composed operator

$$\bar{N}(u) = N'^*_u(u; \psi) \quad \text{with} \quad \psi = K(u; \chi) \quad \text{and} \quad \chi = N(u). \tag{A2.3}$$

The total Gateaux differential is

$$\bar{N}'_u(u; \delta u) = (N'^*_u)'_u(u; \psi, \delta u) + (N'^*_u)'_\psi(u; \delta\psi). \tag{A2.4}$$

To perform the adjoint, we make the scalar product

$$\langle \bar{N}'_u(u, \delta u), \varphi \rangle = \langle (N'^*_u)'_u(u; \psi, \delta u), \varphi \rangle + \langle (N'^*_u)'_\psi(u; \delta\psi), \varphi \rangle. \tag{A2.5}$$

Using the identities (A1.6) and (A1.7), we obtain

$$\langle \bar{N}'_u(u, \delta u), \varphi \rangle = \langle N''_{uu}(u; \varphi, \delta u), \psi \rangle + \langle N'_u(u; \varphi), \delta\psi \rangle. \tag{A2.6}$$

The first scalar product of the second member is symmetric in the couple φ, δu on account of the identity (A1.12). Since

$$\delta\psi = K'_u(u; \chi, \delta u) + K'_\chi(u; \delta\chi) \tag{A2.7}$$

the second scalar product becomes

$$\langle N'_u(u; \varphi), \delta\psi \rangle = \langle N'_u(u; \varphi), K'_u(u; \chi, \delta u) \rangle + \langle N'_u(u; \varphi), K'_\chi(u; \delta\chi) \rangle. \tag{A2.8}$$

From the symmetry condition of K

$$\langle \eta, K(u, \chi) \rangle = \langle \chi, K(u; \eta) \rangle \tag{A2.9}$$

we obtain, by partial differentiation in u

$$\langle \eta, K'_u(u; \chi, \delta u) \rangle = \langle \chi, K'_u(u; \eta, \delta u) \rangle \tag{A2.10}$$

and by partial differentiation in χ

$$\langle \eta, K'_\chi(u; \delta\chi) \rangle = \langle \delta\chi, K(u; \eta) \rangle. \tag{A2.11}$$

Inserting eqns (A2.10) and (A2.11) in eqn (A2.8) we obtain

$$\langle N'_u(u; \varphi), \delta\psi \rangle = \langle N(u), K'_u(u; N'_u(u; \varphi), \delta u \rangle + \langle N'_u(u, \delta u), K(u; N'_u(u; \varphi)) \rangle. \tag{A2.12}$$

The scalar product of the second member is symmetric in φ and δu by condition (A2.1); the second scalar product is symmetric by condition (A2.9). Then the first member of eqn (A2.12) is symmetric: it follows from eqn (A2.6) that

$$\langle \bar{N}'_u(u, \delta u), \varphi \rangle = \langle \bar{N}'_u(u; \varphi), \delta u) \tag{A2.13}$$

which is what we wished to demonstrate.

APPENDIX 3

Gurtin's original method[14] is a particular case of the present method. To prove this, let us observe that Gurtin's method consists of two steps. The first one is a preliminary transformation of the differential equation

into an integro-differential one; the second one is the use of the convolution of two functions. In [46] it was shown that the essential point was the convolution; the preliminary transformation is unnecessary.

In spite of this we want to examine both steps. The transformation of the differential equation into an integro-differential one is performed by a Laplace transform, then by division by s or s^2 (s is the parameter of the Laplace transform) depending on whether the given equation contains first or second-order derivatives, and by an antitransformation. These 3 operations are equivalent to applying the operator

$$Wv(t) = \int_0^T (t-\tau)^{n-1} H(t-\tau) v(\tau)\, d\tau \tag{A3.1}$$

where n is the order of the time derivative.

The second step is equivalent to the application of the convolution operator

$$Cv(t) = v(T-t) = \int_0^T \delta(T-t-\tau) v(\tau)\, d\tau. \tag{A3.2}$$

Combining the two operators, we see that Gurtin's method is equivalent to the use of the integrating operator

$$Rv(t) = CHv(t) = \int_0^T (T-t-\tau)^{n-1} H(T-t-\tau) v(\tau)\, d\tau. \tag{A3.3}$$

As a test, let us show that we obtain one of Gurtin's results. Let us consider the heat-conduction problem ([14], eqn (3.1))

$$\begin{cases} \alpha^2 \nabla^2 u(x,t) - \partial_t u(x,t) = 0 & x \in R \\ u(x,0) = u_0 \quad u(\bar{x},t) = U(\bar{x}) & \bar{x} \in \partial R. \end{cases} \tag{A3.4}$$

The operator is affine and will be denoted as A. To apply the integrating operator $R = CW$ we first apply $W(n=1)$

$$\int_0^T H(t-\tau)[\alpha^2 \nabla^2 u(x,\tau) - \partial_t u(x,\tau)]\, d\tau = \int_0^T \alpha^2 \nabla^2 u(x,\tau)\, d\tau - \int_0^T \partial_t u(x,\tau)\, d\tau$$
$$= \alpha^2 * \nabla^2 u(x,t) - u(x,t) + u_0(x) = 0. \tag{A3.5}$$

which coincides with eqn (3.6) of paper[14]. If we now apply the operator C, the operator becomes potential. Since the application of the convolution operator and the subsequent use of the Cartesian bilinear functional is equivalent to the use of the convolutive bilinear functional

$$\langle v, u \rangle_C = \int_R \int_0^T v(x,t) u(x, T-t)\, dt\, dx \tag{A3.6}$$

the functional will be given by the general formula

$$\bar{F}[u] = \int_0^T \left\langle WA\eta(\lambda), \frac{\partial \eta}{\partial \lambda} \right\rangle d\lambda \tag{A3.7}$$

and letting $\eta(\lambda) = \lambda u(x,t)$, we obtain the functional

$$\bar{F}[u] = \int_R \int_0^T \int_0^1 [\alpha^2 * \nabla^2 \lambda u(x,t) - \lambda u(x,t) + u_0(x)] u(x, T-t)\, d\lambda\, dt\, dx$$
$$= -\frac{1}{2} \int_K \int_0^T [\alpha^2 * \nabla u(x,t) \cdot \nabla u(x, T-t) + u(x,t) u(x, T-t) - 2u_0(x) u(x, T-t)]\, dt\, dx. \tag{A3.8}$$

Since Gurtin considered a functional depending on t, he denoted $\Omega_t(u)$, we obtain it taking a variable t instead of a fixed T. The integral over $(0,t)$ then becomes a convolution product and we obtain

$$\Omega_t(u) = -\frac{1}{2} \int_R [\alpha^2 * \nabla u(x,t) * \nabla u(x,t) + u(x,t) * u(x,t) - 2u_0(x) * u(x,t)]\, dx, \tag{A3.9}$$

which coincides with the functional given by Gurtin ([14], eqn (3.7)].

APPENDIX 4

Magri's method[24] is a particular case of the present formulation. To show this, let us remember that Magri's procedure for giving a variational formulation to linear problems requires two steps. The first one is the definition of a symmetric and non-degenerate bilinear form on $V \times V$, denoted (v_1, v_2); the second one is the definition of a bilinear form on $U \times V$ by the relation $\langle v, u \rangle = (v, Lu)$.

The first bilinear form may be chosen to be of the kind

$$(v_1, v_2) = \int_0^T v_1(t) \int_0^T k(t,\tau) v_2(\tau)\, d\tau\, dt \tag{A4.1}$$

where k is a symmetric positive definite kernel ([24], eqn (3.4)]. The second bilinear form is

$$\langle v, u\rangle = (v, Lu) = \int_0^T v(t) \int_0^T k(t, \tau) Lu(\tau)\, d\tau\, dt \tag{A4.2}$$

and the functional is

$$\bar{F}[u] = \frac{1}{2}\langle Lu, u\rangle = \frac{1}{2}\int_0^T Lu(t) \int_0^T k(t, \tau) Lu(\tau)\, d\tau\, dt. \tag{A4.3}$$

If we denote as

$$\langle v, u\rangle_0 = \int_0^T v(t)u(t)\, dt \tag{A4.4}$$

the Cartesian bilinear form on $U \times V$, the preceding functional may be written as

$$\bar{F}[u] = \frac{1}{2}\langle Lu, KLu\rangle_0 \tag{A4.5}$$

which coincides with the functional given in Theorem 2.

In the two examples given in [24], the operator K is chosen to be of the Volterra kind in order to assure invertibility. The kernels are exactly those given in rows 3 and 4 in Table 1. Since the corresponding operators K are not positive definite, the functionals are not minimal but only stationary. Any other kernel from Table 1 would have caused Magri's functionals to be extrema.

APPENDIX 5

Telega[39] gives a first extension of Magri's rule to nonlinear operators. The idea may be summarized as follows.

Starting with a symmetric, non-degenerate bilinear functional (v_1, v_2) on $V \times V$ and taking an element u_0, since $N'_u(u_0; \varphi)$ is linear on φ one may define a bilinear functional on $V \times U$ as

$$\langle v, \varphi\rangle_0 = (v, N'_u(u_0, \varphi)). \tag{A5.1}$$

now

$$\langle N'_u(u; \varphi), \psi\rangle_0 = (N'_u(u; \varphi), N'_u(u_0; \psi)) \tag{A5.2}$$

and

$$\langle N'_u(u; \psi), \varphi\rangle_0 = (N'_u(u; \psi), N'_u(u_0; \varphi)). \tag{A5.3}$$

If the operator N is such that an element u_0 exists for which the condition

$$(N'_u(u; \varphi), N'_u(u_0; \psi)) = (N'_u(u; \psi), N'_u(u_0; \varphi)) \tag{A5.4}$$

holds, it follows that

$$\langle N'_u(u; \psi), \varphi\rangle_0 = \langle N'_u(u; \varphi), \psi\rangle_0 \tag{A5.5}$$

and thus the operator N'_u is symmetric with regard to the bilinear functional (A5.1): it follows that it admits of a variational formulation.

We see that the variational formulation is possible for those operators that satisfy condition (A5.4).