

On the variational formulation for linear initial value problems.

ENZO TONTI (Milano) (*) (**)

Summary. – *The paper shows the key place of the choice of the bilinear form in order to give a variational formulation to a given problem. In particular it is shown how the use of a convolution bilinear form makes possible a variational formulation for linear initial value problems.*

A critical survey of the three main methods that was devised in the past to solve the same problem is done.

Sunto. – *La nota mette in evidenza il ruolo fondamentale della scelta di una forma bilineare al fine di dare formulazione variazionale ad un dato problema. In particolare è mostrato come l'uso di una forma bilineare di convoluzione renda possibile la formulazione variazionale dei problemi ai valori iniziali.*

Si fa un esame critico dei tre principali metodi che sono stati escogitati nel passato per risolvere lo stesso problema.

1. – Introduction.

A variational formulation for initial value problems is not possible in the classical context of the calculus of variations. The reason lies in that we persist to use as scalar product that given by

$$(1.1) \quad \int_0^T u(t)v(t) dt$$

or one of its variants (with weight factors, extension to more variables, etc.). It is easily seen that with this scalar product initial value problems do not fit the necessary conditions in order that a variational formulation does exist.

Things change when we decide to abandon the scalar product (1.1) and to introduce other bilinear forms as, for ex. that given by the convolution of two functions

$$(1.2) \quad \int_0^T u(T-t)v(t) dt .$$

With this bilinear form, in particular, we can do variational formulation to initial value problems.

(*) This work has been sponsored by the Consiglio Nazionale delle Ricerche. Author address: Istituto di Matematica del Politecnico, Piazza Leonardo da Vinci, 32, 20133 Milano (Italia).

(**) Entrata in Redazione il 19 novembre 1971.

Then the theorem asserting that in order that an equation be the Euler-Lagrange equation for some functional it is necessary that its operator be symmetric (if linear) or its Gateaux derivative be symmetric (if nonlinear), must not be considered so severe as it has been up to now. If an operator is not symmetric with a given bilinear form it can become symmetric with *another* bilinear form. *Then the choice of different bilinear forms becomes the key to try to give variational formulation to new class of equations.*

In § 2 we give a brief account of the ideas and tools for a modern foundation of calculus of variations stressing the role of bilinear forms, adjoint of an operator, symmetric operators, circulation of an operator along a line in the function space.

In § 3 we examine the various methods devised to give variational formulation to initial value problems.

In § 4 a convolution bilinear form is introduced and it is shown that linear operators that describe initial value problems become symmetric with such bilinear form. Then a variational formulation becomes possible.

In § 5 we make a comparison of the present method with other existing methods.

2. - The essential of calculus of variations in operator form.

2.1. Bilinear forms.

Given two vector spaces U and V , with the name *bilinear form* or *bilinear functional* is meant a functional $F[u, v]$ that is linear in u and linear in v with $u \in U$, $v \in V$. Once a bilinear functional is introduced we say that the two vector spaces are put in *duality* [17, p. 88].

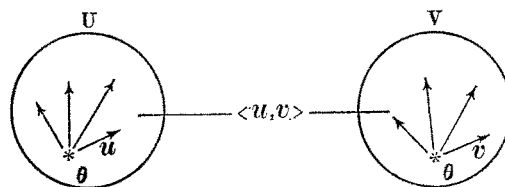


Fig. 1.

The duality is called *separating* in v if, given an arbitrary $u \in U$ with $u \neq \theta$, it exist at least one v such that $\langle u, v \rangle \neq 0$ [17, p. 88]. The reason for the request that the duality be separating lies in the fact that if $\langle u, v \rangle = 0$ for *every* u that belongs to a dense subset of U then must be necessarily $v = \theta$. In the sequel we shall consider only dualities that are separating with respect to u and v .

To choose a topology for both spaces the natural request is that they be such that for *every* linear functional $l[u]$ an element v_0 can be found such that $l[u] = \langle u, v_0 \rangle$

and analogous requirement on V . Last request is the analogous of Riesz representation theorem. Such topology will be said *compatible* with the duality [18, p. 32] [19, p. 152].

REMARK I. – In books of functional analysis the concept of dual space of a given vector space is more frequent than that of spaces in duality. The *algebraic dual* of a linear space is the space of all *linear* functionals defined on that space; and the *topological dual* of a topological linear space is the space of all *linear and continuous* functionals defined on that space. Now with the advent of the theory of spaces in duality, one of the great merits of Bourbaki group [17, p. 88], the concept of dual space can be absorbed in that theory. From the physical point of view the idea of putting two spaces in duality is more basic than that of dual space. So in mechanics we learn displacement, forces and the bilinear form that gives the work $\langle f, s \rangle = \sum_k f_k s_k$ in this order. This makes useless to conceive a force as a « linear functional » or « linear form » on displacements.

2.2 *The adjoint of an operator.*

a) Let us consider a linear operator L from a linear space U to a linear space U' and two more vector spaces V and V' that we put in separating duality with U and U' respectively, by means of *two* bilinear forms

$$(2.1) \quad \langle u, v \rangle_I, \quad \langle u', v' \rangle_{II}.$$

For the two pairs of linear spaces in duality we choose topologies that are compatible with the dualities (see for ex. § 4.2).

The adjoint of the linear operator L is by definition, a linear operator \tilde{L} from V' to V that satisfies the identity

$$(2.2) \quad \langle Lu, v' \rangle_{II} \equiv \langle u, \tilde{L}v' \rangle_I$$

for every $u \in U$. The adjoint \tilde{L} is a mapping from V' to V i.e. $\tilde{L}v' = v$ (see fig. 2).

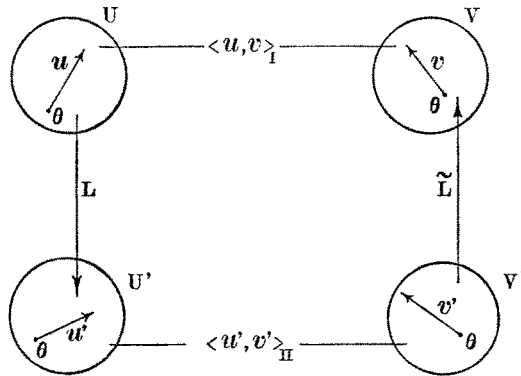


Fig. 2.

We stress the fact that to define the adjoint of an operator we need *two* bilinear forms and *four* spaces.

b) A particular case arises when we have a linear mapping between two spaces that are put in duality by the bilinear form $\langle u, v \rangle$. In such case the adjoint of the operator L satisfies the relation

$$(2.3) \quad \langle u_2, Lu_1 \rangle = \langle \tilde{L}u_2, u_1 \rangle$$

and the corresponding diagram is shown in fig. 3.

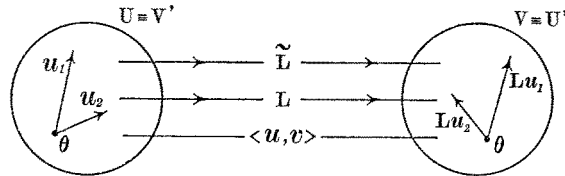


Fig. 3.

c) Another particular case arises when U and V are two Hilbert spaces and the bilinear forms are the scalar products in the two spaces, $(u_1, u_2)_I$ and $(v_1, v_2)_{II}$ (fig. 4). The adjoint of L satisfies the relation

$$(2.4) \quad (Lu, v)_{II} = (u, \tilde{L}v)_I$$

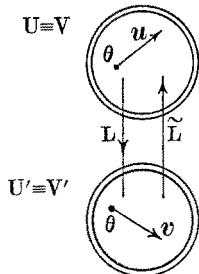


Fig. 4.

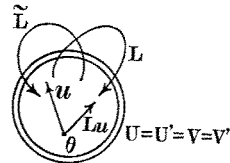


Fig. 5.

d) A last particular case arises when U is an Hilbert space with scalar product (u_1, u_2) and L go from U to U (fig. 5) then the definition of adjoint (2.2) reduces to

$$(2.5) \quad (Lu_1, u_2) = (u_1, \tilde{L}u_2).$$

This is the most common definition of adjoint of a linear operator [20, p. 43].

Definition given in *a*) and the three particular cases *b*), *c*), *d*) require that the domain of L be all the linear space U . Differential operators defined in L_2 -spaces (more in general in L_p -spaces) cannot have as domain the whole space but a dense subset $D(L)$ of U . Moreover they are not continuous in such spaces. In this case the problem arises to define the domain $D(\tilde{L})$ of the adjoint operator \tilde{L} . By definition the domain of \tilde{L} is the set of elements v' such that the functional $\langle Lu, v' \rangle$ be continuous in spite of the discontinuous character of L .

2.3. *Symmetric and selfadjoint operators.*

A linear operator is called *symmetric* with respect to a bilinear form $\langle u, v \rangle$ if it satisfies the condition

$$(2.6) \quad S \subseteq \tilde{S}$$

i.e. if \tilde{S} is an extension of the operator S : this means that $D(\tilde{S}) \supset D(S)$ and that $\tilde{S}u = Su$ for $u \in D(S)$.

An operator is called *selfadjoint*, with reference to a bilinear form, if $S = \tilde{S}$.

The symmetry of an operator can occur only when the operator works between two spaces in duality i.e. in the case *b*) and *d*) of § 2.2.

An operator is called *formally selfadjoint* when the corresponding formal differential operator is equal to its adjoint [30, p. 1287].

Among the properties of symmetric operators we shall use the following:

a) The square of a symmetric operator is also symmetric.

PROOF. – From $S \subseteq \tilde{S}$ follows for every $u, v \in D(S)$

$$(2.7) \quad \langle v, S^2u \rangle = \langle v, S\tilde{S}u \rangle = \langle Sv, Su \rangle = \langle S^2v, u \rangle .$$

b) The inverse of a symmetric operator is also symmetric.

PROOF. – From $\langle v, Su \rangle = \langle Sv, u \rangle$ putting $u = S^{-1}f$, $v = S^{-1}g$ follows $\langle S^{-1}g, f \rangle = \langle g, S^{-1}f \rangle$.

c) If S is linear and symmetric and u_0 is one solution of the problem $Su = f$ then u_0 is critical point of the functional

$$(2.8) \quad F[u] = \frac{1}{2} \langle u, Su \rangle - \langle u, f \rangle$$

i.e. $\delta F[u]|_{u_0} = 0$ and viceversa [21, p. 75].

d) If the symmetric operator S is positive i.e. $\langle u, Su \rangle \geq 0$ then the functional (2.8) is minimum at u_0 i.e. $\delta^2 F[u]|_{u_0} \geq 0$ [21, p. 75].

2.4. *Conservative vector fields.*

Calculus of variations requires as a fundamental tool a *bilinear form*, i.e. a bilinear functional. To show this we can start with the discrete form of the calculus of variations, i.e. *the theory of conservative vector fields in finite dimensional spaces*. Nothing is lost in generality and much is gained in clarity if we limit ourselves to fields in the ordinary three dimensional spaces \mathbf{R}^3 .

Let us consider a vector field $v = v(u)$ where u is the radius vector, and be $\Omega \subset \mathbf{R}^3$ a convex ⁽¹⁾ region in which the vector field is defined. In order to introduce the concept of conservative vector field we must define the *circulation* of a vector along a line and this requires the introduction of a bilinear form as

$$(2.9) \quad \langle u, v \rangle = v_x u_x + v_y u_y + v_z u_z.$$

A line is described by a variable vector $\eta = \eta(\lambda)$ where λ a parameter that we choose for convenience so that $\lambda = 0$ for $\eta = u_0$ and $\lambda = 1$ for $\eta = u_1$

$$(2.10) \quad C = \int_{\lambda=0}^{\lambda=1} \langle d\eta(\lambda), v[\eta(\lambda)] \rangle.$$

The circulation depends, in general, from the line connecting the two points u_0 and u_1 : when the circulation between two arbitrary points does not depend from the line connecting them we say that the vector field is *conservative*.

Choosing a fixed element u_0 in the region in which the field is defined, we can consider the circulation from u_0 to whatever u in the region and this circulation depends only on u (for fixed u_0). In this way a number can be associated to every u and we have constructed a *scalar field*. The function $f(u)$ so defined is called the *potential* of the vector field. Since the region Ω is convex we can choose straight lines to calculate the circulation from u_0 to u . Then $\eta(\lambda) = u_0 + \lambda(u - u_0)$ and

$$(2.11) \quad d\eta(\lambda) = \frac{d\eta}{d\lambda} d\lambda = (u - u_0) d\lambda.$$

Relation (2.10) then gives

$$(2.12) \quad f(u) = f(u_0) + \int_{\lambda=0}^{\lambda=1} \langle (u - u_0), v[u_0 + \lambda(u - u_0)] \rangle d\lambda.$$

⁽¹⁾ A convex region is such that for every pair of points contained in it all points of the segment of straight line connecting them belong to the region. Then a convex region is simply connected. Particular convex regions are the affine manifolds: in \mathbf{R}^3 these are planes that does not contain the origin.

In order that a vector field be conservative in the region Ω the circulation must vanish along any closed path. In particular this must be true for infinitesimal closed paths as the one indicated in fig. 6.

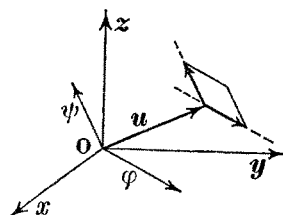


Fig. 6.

Simple calculations [2] show that this path-independence is equivalent to the requirement that the vector field satisfies the condition

$$(2.13) \quad \sum_{h,k} \varphi_h a_{hk} \psi_k = \sum_{h,k} \psi_h a_{hk} \varphi_k$$

where

$$(2.14) \quad a_{hk} \stackrel{\text{def}}{=} v_{h,k}.$$

Condition (2.13) amount to the request of symmetry of the matrix a_{hk} formed with the partial derivatives (2.14). If the vector field satisfies condition (2.13) at every point of the region Ω in which the field is defined, and the region Ω is simply connected then the circulation does not depends from the path for every line contained in the region Ω .

Note that the condition of symmetry of a_{hk} implies that

$$(2.15) \quad v_{h,k} - v_{k,h} = 0$$

that is the statement that the curl of the vector field vanishes. Using an operator notation relation (2.13) can be written as (see [2])

$$(2.16) \quad \langle \varphi, v'_u \psi \rangle = \langle \psi, v'_u \varphi \rangle.$$

A fundamental property of conservative vector fields (that becomes the key of the calculus of variations) is that the points in which the vector field vanishes are the points in which the potential is stationary and viceversa $v(u_0) = 0 \Leftrightarrow \delta f(u_0) = 0$ [2, p. 143]. These are called *critical points* [1, p. 77].

REMARK II. — An essential point that is not commonly stressed in books on calculus of variations is that the circulation depends from the bilinear form chosen. Then if we choose *another* bilinear form as for ex.

$$(2.17) \quad \langle u, v \rangle_c = u_x v_x + u_y v_y + u_z v_z$$

we can define a new circulation with this bilinear form; then we can expect that some vector field that was not conservative with respect to the old bilinear form (2.9) becomes conservative with respect the new bilinear form (2.17). This indicates that *the fact that a vector field be conservative is not an intrinsic property of the vector field but depends from the bilinear form used.*

It is autor's feeling that this remark may becomes the starting point for an indefinite enlargement of the field of application of the calculus of variations in future.

These considerations extend from spaces of finite dimensions to *functions spaces* where a nonlinear mapping $v = N(u)$ can be considered as describing a vector field on a function space U . The domain of the nonlinear operator N will be assumed to be a *convex* region $D(N)$ of the function U -space.

In this case the bilinear form corresponding to (2.9) is

$$(2.18) \quad \langle u, v \rangle = \int_{\Omega} u(P)v(P) d\Omega .$$

A line in a function space becomes a one parameter family of functions $\eta(P, \lambda)$. The circulation becomes

$$(2.19) \quad C = \int_{\lambda=0}^{\lambda=1} \langle \delta\eta(P; \lambda), N[\eta(P; \lambda)] \rangle$$

where

$$(2.20) \quad \delta\eta(P, \lambda) = \frac{\partial\eta}{\partial\lambda} d\lambda .$$

A stright line from $u_0(P)$ to $u(P)$ is given by

$$(2.21) \quad \eta(P, \lambda) = u_0(P) + \lambda[u(P) - u_0(P)]$$

and the potential is [1, p. 58]

$$(2.22) \quad F[u] = F[u_0] + \int_{\lambda=0}^{\lambda=1} \langle u(P) - u_0(P), N\{u_0(P) + \lambda[u(P) - u_0(P)]\} \rangle d\lambda .$$

If the domain of the operator $N(u)$ contains the null element θ (this surely happens if the operator is linear) then we can simplify the formula choosing $u_0(P) = \theta$.

Then

$$(2.23) \quad F[u] = F[\theta] + \int_{\lambda=0}^{\lambda=1} \langle u(P), N[\lambda u(P)] \rangle d\lambda .$$

The condition of path-independence is reduced to the requirement that [1, p. 59]

$$(2.24) \quad \langle \psi, N'_u \varphi \rangle = \langle \varphi, N'_u \psi \rangle$$

that states the symmetry of the operator N'_u defined by

$$(2.25) \quad \lim_{\varepsilon \rightarrow 0} \frac{N(u + \varepsilon \varphi) - N(u)}{\varepsilon} = N'_u \varphi$$

and called *Gateaux derivative* of the nonlinear operator N [1, p. 56].

The fulfilment of relation (2.24) is equivalent to the statement of path-independence of the circulation for whatever line in function space if the domain of the operator is simply connected [1, p. 32]: since we have chosen a convex domain this requirement is fulfilled.

REMARK III. — Once more we observe that the fulfilment of relation (2.24) depends basically from the bilinear form chosen. Then if we choose a bilinear form different from (2.18) we can hope that operators that are not of potential kind with that bilinear form become potential now.

Then when we have an operator that does not satisfy the symmetry condition (2.24) with a given bilinear form one of the first possibility to be explored is the change of the bilinear form. Phrases as «the adjoint of this operator is that operator» have a meaning only for a specified bilinear form.

In sec. 4 we present a kind of bilinear form that makes symmetric first order differential operators; this opens the way to give variational formulation to initial value problems, usually excluded from variational treatments.

Before this we give a survey of the three main methods devised to give variational formulation to initial value problems.

3. — Survey of existing methods.

Various attempts were made in the past and in recent years to give a variational formulation to initial value problems using different methods. The most used of them are:

- a) a method for formally selfadjoint operators;
- b) the method of adding the adjoint equation;

e) Gurtin method of convolution.

We summarize briefly these methods in order to compare them critically ⁽²⁾.

3.1. *The method for formally selfadjoint operators.*

This is historically the first of the tricks devised to overcome the difficulty mentioned. Let the initial value problem be described by a linear operator that is *formally selfadjoint* with the usual bilinear form

$$(3.1) \quad \langle u, v \rangle = \int_0^T u(t)v(t)dt .$$

This is the case of the operator of the following problem

$$(3.2) \quad \begin{cases} \frac{d^2}{dt^2} u(t) = f(t) \\ u(0) = 0 \\ u'(0) = 0 \end{cases} \quad \text{whose adjoint is} \quad \begin{cases} \frac{d^2}{dt^2} v(t) = g(t) \\ v(T) = 0 \\ v'(T) = 0 . \end{cases}$$

The operator is formally selfadjoint because its formal part, i.e. d^2/dt^2 is equal to the formal part of its adjoint, but the domain of the operator $D(L)$ is neither equal nor contained in that of its adjoint, and then it is not symmetric (see § 2.3).

The method from long time devised is that of ignoring the initial condition $u'(0) = 0$ and add an artificial condition $u(T) = 0$. In this way the original *initial* value problem is transformed into a *boundary* value problem that, this time, is governed by a *symmetric* operator

$$(3.3) \quad \begin{cases} \frac{d^2}{dt^2} u(t) = f(t) \\ u(0) = 0 \\ u(T) = 0 \end{cases} \quad \text{whose adjoint is (with} \\ \text{the bilinear form (3.1))} \quad \begin{cases} \frac{d^2}{dt^2} v(t) = g(t) \\ v(0) = 0 \\ v(T) = 0 . \end{cases}$$

Of course the solution of the problem (3.3) is not a priori solution of the problem (3.2) ⁽³⁾.

Problem (3.3) admits a variational formulation as follows: the solution of the

⁽²⁾ In this critical survey we omit the necessary details on the functional class of functions and on the kinds of function spaces: these matters will be examined in following sections.

⁽³⁾ For example if $f(t) = 1$ and $0 < t < 1$ the unique solution of problem (3.2) is $u(t) = \frac{1}{2}t^2$ while the unique solution of the corresponding problem (3.3) is $u(t) = \frac{1}{2}t(t-1)$.

boundary value problem (3.3) makes stationary the functional

$$(3.4) \quad F[u] = \int_0^T \left\{ \frac{1}{2} [u'(t)]^2 - f(t)u(t) \right\} dt$$

among all functions $u(t)$ that satisfy the same boundary conditions of problem (3.3).

The unique advantage of this variational formulation lies in that the Euler-Lagrange equation of the functional (3.4) is the differential equation of the problem (3.2).

But on account of the artificial change from an initial to a final condition direct methods cannot be applied to find the solution of the ordinary problem (3.2), and existence or uniqueness of the solution of (3.2) cannot be inferred from the functional.

Briefly stated: *the functional (3.4) does not give a variational formulation to the initial value problem (3.2) but to the boundary value problem (3.3) artificially associated to it.*

This procedure extends also to nonlinear operators whose Gateaux derivative (see formula 2.25) is formally selfadjoint with the bilinear form (3.1).

Even with this handicap this technique is much used in mechanics and theoretical physics: Hamilton principle, the stationary action principles of electromagnetism, quantum mechanics, general relativity, etc. rest upon this artificial transformation of an initial to a boundary value problem ⁽⁴⁾.

3.2. *The method of the adjoint equation.*

The previous technique of changing initial to boundary value problem cannot be applied when the operator is not at least formally selfadjoint. This is the case of the Fourier equation of heat conduction that contains first order partial time derivatives. This is also the case of many dissipative systems characterized by irreversible phenomena.

One method goes back to MORSE-FESHBACH [4, p. 298] and consists in the considering alongside with the ordinary differential equation another artificial equation with the adjoint operator (then with final conditions). So for Fourier's equation

$$(3.5) \quad \begin{cases} \Delta\theta(P, t) - k \frac{\partial\theta(P, t)}{\partial t} = f(P, t) \\ \theta(P, 0) = 0 \\ \theta(P, t)|_S = 0 \end{cases} \quad \rightarrow \quad \begin{cases} \Delta\varphi(P, t) + k \frac{\partial\varphi(P, t)}{\partial t} = g(P, t) \\ \varphi(P, T) = 0 \\ \varphi(P, t)|_S = 0 \end{cases}$$

where P belong to a region Ω whose boundary we indicate with S and $0 \leq t \leq T$.

⁽⁴⁾ We remember that Hamilton principle states that the natural motion is privileged among those conceivable motions that have the same *initial* and *final* configuration, then ignoring the physical initial conditions on momenta.

More in general

$$(3.6) \quad \begin{cases} Lu = f & \text{be the given equation} \\ \tilde{L}v = g & \text{is the adjoint equation.} \end{cases}$$

These two equations can be cast in the form of the matrix-differential equation

$$(3.7) \quad \begin{pmatrix} 0 & \tilde{L} \\ L & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} g \\ f \end{pmatrix}$$

whose operator is symmetric [5, p. 116], and then a variational formulation becomes possible. The functional can be found with a general procedure [2, p. 159] obtaining

$$(3.8) \quad F[u, v] = \langle v, Lu \rangle - \langle f, v \rangle - \langle g, u \rangle$$

About this expedient Morse-Feshbach say «By this arbitrary trick we are able to handle dissipative systems as though they were conservative. This is not very satisfactory if an alternate method of solution is known, but it will be necessary, in order to make any progress when we come to study dissipative fields, as in the diffusion equation». [4, p. 299].

From the physical point of view the great problem is to justify the «mirror» equation. Which meaning can be attached to the adjoint function $\varphi(P, t)$, to the function $g(P, t)$, to the final condition $\varphi(P, T) = 0$?

Various attempts were made in specific fields where this expedient is used, as in the domain of reactor physics for transport equation [6, p. 85], [7] to find possible meanings of these adjoint quantities. Apart from some local success the method reveals its character of mathematical trick.

For applications of this method see [23], [8], [9], [25]. The troubles of such method are discussed in detail in [24, p. 245].

3.3. Gurtin method of convolution.

In 1964 GURTIN [10] showed how the convolution of two function $u(t)$ and $v(t)$

$$(3.9) \quad z(t) \stackrel{\text{def}}{=} \int_0^t u(t-\tau)v(\tau)d\tau$$

permits to give a variational formulation to the linear initial value problem. Gurtin procedure is very sophisticated: we summarize here the fundamental ideas. Let us consider for instance the wave equation

$$(3.10) \quad \begin{cases} \Delta\varphi(P, t) - \frac{1}{c^2} \frac{\partial^2\varphi(P, t)}{\partial t^2} = 0 \\ \varphi(P, 0) = \alpha(P) \\ \left. \frac{\partial\varphi(P, t)}{\partial t} \right|_{t=0} = \beta(P) \\ \varphi(Q, t) = 0 \end{cases}$$

where for simplicity we consider homogeneous boundary conditions. The first step is to replace the given problem with an equivalent one. To this end a preliminary Laplace transform is done on the function $\varphi(P, t)$. Putting

$$(3.11) \quad \Phi(P, s) \stackrel{\text{def}}{=} \int_0^{\infty} \exp(-st) \varphi(P, t) dt = \mathfrak{L}[\varphi]$$

equation (3.10) becomes

$$(3.12) \quad \Delta \Phi(P, s) - \frac{1}{c^2} s^2 \Phi(P, s) = -\frac{1}{c^2} [s\alpha(P) + \beta(P)].$$

Now we divide the equation by s^2 obtaining

$$(3.13) \quad \frac{1}{s^2} \Delta \Phi(P, s) - \frac{1}{c^2} \Phi(P, s) = -\frac{1}{c^2} \frac{1}{s} \alpha(P) - \frac{1}{c^2} \frac{1}{s^2} \beta(P).$$

At this point if we take the inverse transform $\mathfrak{L}^{-1}[\Phi]$ we no longer obtain derivatives in the second term because of the disappearance of the function s^2 in the second term. Because the inverse transform of the product of the two functions $1/s^2$ and $\Delta \Phi(P, s)$ is equal to the convolution of the two primitives according to the rule

$$(3.14) \quad \mathfrak{L} \left[\int_0^t u(t-\tau) v(\tau) d\tau \right] \equiv \mathfrak{L}[u(t)] \cdot \mathfrak{L}[v(t)]$$

then the inverse Laplace transform applied to eq. (3.13) gives

$$(3.15) \quad \int_0^t (t-\tau) \Delta \varphi(P, \tau) d\tau - \frac{1}{c^2} \varphi(P, t) = -\frac{1}{c^2} \beta(P) - \frac{1}{c^2} t\alpha(P).$$

The equation obtained is now integro-differential ⁽⁵⁾. The functions $\varphi(P, t)$ that

⁽⁵⁾ Accidentally we observe that equation (3.15) can be obtained more easily without the use of Laplace transform as follows: writing eq. (3.10) as

$$\frac{\partial^2 \varphi(P, t)}{\partial t^2} = c^2 \Delta \varphi(P, t)$$

and integrating twice on t we obtain

$$\varphi(P, t) = c^2 \int_0^t (t-\tau) \Delta \varphi(P, \tau) d\tau + A(P)t + B(P)$$

where use was made of Dirichlet relation

$$\int_0^t \int_0^{\tau} f(\tau) d\tau dt = \int_0^t (t-\tau) f(\tau) d\tau.$$

Imposing the initial conditions given we find $A(P) = \alpha(P)$, $B(P) = \beta(P)$ eq. (3.15) is thus obtained.

satisfy this equation *automatically* satisfy the two initial conditions in (3.10). Then the given problem is equivalent to the problem: solve the eq. (3.15) with the boundary condition $\varphi(S, t) = 0$.

Now we look for a variational formulation of eq. (3.15). The convolution kernel in eq. (3.15) of the integro-differential equation is not symmetric with respect to the ordinary bilinear form (3.1). But it becomes symmetric with respect to the bilinear map

$$(3.16) \quad \iiint_V \int_0^t \varphi(P, t - \tau) \psi(P, \tau) d\tau dV$$

then, as can be shown, we can deduce eq. (3.15) by the stationarity of $A_t[\varphi]$ given by

$$(3.17) \quad A_t[\varphi] = \iiint_V \int_0^t \varphi(P, t - \tau) [\varphi(P, \tau) - 2\alpha(P) - 2\tau\beta(P)] + \\ + \iiint_V \int_0^t \left\{ e^{\lambda(t-\tau)} \int_0^\tau \nabla\varphi(P, \tau - \sigma) \cdot \nabla\varphi(P, \sigma) d\sigma \right\} d\tau dV.$$

$A_t[\varphi]$ is, for fixed t , a functional of φ .

The stationarity of the functional with respect to the functions $\varphi(P, t)$ that satisfy the boundary condition $\varphi(S, t) = 0$ lead to eq. (3.15) as Euler-Lagrange equation.

This procedure was used extensively by Gurtin himself [10], [11], [12] and others [13], [14], [15], [16]. Functionals obtained by this method are often complicate. Nevertheless the method has on the previous ones the merit that neither artificial boundary conditions (as in the method *a*) nor artificial differential equations (as in the method *b*) are added. Only manipulations on the given equations are done. The functions that make stationary the functional are exactly those that satisfy a problem *equivalent* to the one given at the beginning.

This fact indicates that convolution is a natural tool to treat initial value problems. Previous Laplace transforming is not necessary as shown in footnote (5) and then it has not essential role in the method. It becomes natural to try to simplify this method to obtain directly the given equation.

A progress in this direction arises when we realize that a variational formulation requires the symmetry for the operator and this in turn is dependent on the kind of bilinear form. Then if we want to give a variational formulation to initial value problems we must find a bilinear form, if any, that makes symmetric the corresponding operator. A bilinear form using convolution of two functions gives the hope to be what we need. As a matter of fact we show in this paper that this is the case.

In order to give a detailed and comprehensive exposition of the new method we summarize some useful concepts.

4. - The convolution bilinear form.

4.1. *The convolution product of two vectors.*

Given two finitely dimensional vector spaces U and V , both of the kind \mathbf{R}^n , an usual bilinear form on the elements of the two spaces is

$$(4.1) \quad \langle u, v \rangle = u_1 v_1 + u_2 v_2 + \dots + u_{n-1} v_{n-1} + u_n v_n .$$

We shall introduce another bilinear form, that we call *convolution of the two vectors*

$$(4.2) \quad \langle u, v \rangle_c \stackrel{\text{def}}{=} u_1 v_n + u_2 v_{n-1} + \dots + u_{n-1} v_2 + u_n v_1 .$$

With such bilinear form the two spaces U and V are put in duality. We show that this is a separating duality. In fact if $\langle u, v \rangle_c = 0$ for every u , this implies that $v = \theta$ as can be seen considering for u successively the vectors $(1, 0, 0 \dots 0)$, $(0, 1, 0, 0, \dots 0)$, $\dots (0, 0, \dots 0, 1)$.

Now a matrix that is symmetric with respect to the secondary diagonal, like, for ex.

$$(4.3) \quad L = \begin{pmatrix} a & m & p & f \\ q & b & e & p \\ s & d & b & m \\ c & s & q & a \end{pmatrix}$$

is symmetric with respect to the convolution bilinear form.

Particular cases of such matrices are the following

$$(4.4) \quad \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

that are the finite difference forms associated with the differential operators d/dx and d^2/dx^2 respectively.

This shows in a very incisive way that the concept of symmetry of a matrix depends strongly from the bilinear form we choose. Usually a real matrix is called symmetric if the entries that are symmetrically disposed with respect to the main diagonal are equal: this definition is based on the use of the bilinear form (4.1).

As consequence of the example shown we are convinced that *the notion of symmetry for a matrix is not an priori property but has a meaning only with reference to a prescribed bilinear form*. In particular it cannot be based on the symmetry of the entries of the matrix with respect to interchanging rows with columns.

A relevant difference between the two bilinear forms arises when we try to construct with them a *scalar product* (u, v) between two elements of the *same* space. In order that this be possible, two additional properties must be satisfied:

$$(4.5) \quad (u, v) = (v, u)$$

$$(4.6) \quad (u, u) \geq 0 \quad \text{and} \quad (u, u) = 0 \Rightarrow u = \theta$$

Now the first of these properties is satisfied by both bilinear forms: in the case of convolution bilinear form the property follows from the identity

$$(4.7) \quad \sum_1^n u_k v_{n-k+1} \equiv \sum_1^n u_{n-k+1} v_k.$$

On the contrary the second property is satisfied by the first bilinear form, as is well known, while it is not satisfied by the second bilinear form. To show this a counter-example is sufficient: let $u = (1, 2, -3)$ then

$$(4.8) \quad \langle u, u \rangle_c = (1)(-3) + (2)(2) + (-3)(1) = -2 < 0.$$

Then the convolution product of two vectors cannot be taken as a scalar product of two elements of the same space, i.e. with a convolution bilinear form we cannot construct an Hilbert space⁽⁶⁾. Moreover we cannot construct a Banach space.

4.2. *The convolution bilinear form for functions.*

The analogous with functions of the two bilinear forms (4.1) and (4.2) are respectively

$$(4.9) \quad \langle u, v \rangle = \int_0^x u(t)v(t)dt$$

⁽⁶⁾ Spaces with a scalar product that does not satisfy condition (4.6) are of some interest in physics particularly in relativistic quantum mechanics. For the mathematical literature see [27], [28], [1, p. 89]. For applications in physics see [29].

$$(4.10) \quad \langle u, v \rangle_c = \int_0^T u(T-t)v(t) dt.$$

If U and V are two vector spaces and $u \in U, v \in V$, then everyone of these bilinear forms put the two spaces in duality. The duality is separating as can easily be shown. A topology compatible with the duality for both spaces is $U = L_p[0, T]$ and $V = L_q[0, T]$ with $1/p + 1/q = 1$. This choice assures that the bilinear forms are both continuous in u and v and moreover that every linear functional in u (or in v) can be represented by means of the bilinear forms.

Let us consider the operator

$$(4.11) \quad L \begin{cases} \frac{d}{dt} \\ 0 \leq t \leq T \\ u(0) = 0 \\ u(t) \in C^1[0, T]. \end{cases}$$

We search its adjoint with the scalar products (4.9) and (4.10) respectively. Is

$$(4.12) \quad \langle u, Lv \rangle = \int_0^T u(t)v'(t) dt = \int_0^T -u'(t)v(t) dt + [u(T)v(T) - u(0)v(0)]$$

$$(4.13) \quad \langle u, Lv \rangle_c = \int_0^T u(T-t) \frac{dv(t)}{dt} dt = \int_0^T \frac{d}{dt} u(t)v(T-t) dt + [u(0)v(T) - u(T)v(0)].$$

Then the adjoints of the operator L are respectively

$$(4.14) \quad \tilde{L} = \begin{cases} -\frac{d}{dt} \\ v(T) = 0 \\ v(t) \in \text{A.C.} \end{cases} \quad \tilde{L} = \begin{cases} \frac{d}{dt} \\ v(0) = 0 \\ v(t) \in \text{A.C.} \end{cases}$$

with $\langle u, v \rangle$ with $\langle u, v \rangle_c$

where A.C. denotes the class of absolutely continuous functions. Then while in the first case the operator L is not symmetric, in the second case it is, since $L \subset \tilde{L}$ (?). Thus we discover that *the convolution bilinear form makes symmetric the operator d/dt with homogeneous initial conditions*. Since the symmetry is necessary

(?) The inclusion sign depends from the fact that functions of class $C^1[0, T]$ form a subset of functions A.C. $[0, T]$.

condition in order that a linear operator be the gradient of a functional and it becomes sufficient if the domain is simply connected [1, p. 32], it follows that in this way one can give a variational formulation to linear initial value problems [22].

The functional is found with formula (2.23): see § 4.1. Let us consider the operator

$$(4.15) \quad Q = \begin{cases} \frac{d^2}{dt^2} \\ u(0) = 0 & 0 \leq t \leq T \\ u'(0) = 0 & u(t) \in C^2[0, T] \end{cases}$$

that appears in initial value problems for second order ordinary differential equations. Since this operator is the square of the symmetric operator L given in (4.11), i.e. $Q = L^2$ it follows that it is also a symmetric operator (see § 2.3)

REMARK IV. — The operators L and Q are not positive on account of the indefinite character of the convolution bilinear form: then the variational formulation leads to critical points that are not minima. Since the inverse of a symmetric operator is also symmetric (see § 2.3) and the inverse of a differential operator is an integral operator we are led to examine which kind of integral operators arises by inversion of operator (4.15). Now if

$$(4.17) \quad \frac{d^2}{dt^2} u(t) = f(t), \quad u(0) = 0, \quad u'(0) = 0$$

integrating twice by parts we obtain

$$(4.18) \quad u(t) = \int_0^t (t - \tau) f(\tau) d\tau.$$

The kernel of this integral operator is of the kind known as « convolution » kernel and then this operator is symmetric with respect to the convolution bilinear form (4.10). More in general we show in the next paragraph that all integral operators with convolution kernel are symmetric with the convolution bilinear form.

4.3. *The symmetry of Fredholm and Volterra operators with convolution kernel.*

We now show that the Fredholm integral operator

$$(4.19) \quad \mathcal{F} = \int_0^x k(t - \tau) \dots d\tau$$

that is not symmetric with respect to the bilinear form $\langle u, v \rangle$ becomes symmetric

with respect to the convolution bilinear form $\langle u, v \rangle_c$ given by (4.10)

$$(4.20) \quad \langle u, \mathcal{F}v \rangle_c \equiv \int_0^T \left[u(T-t) \int_0^T k(t-\tau) v(\tau) d\tau \right] dt \equiv \\ \equiv \int_0^T \int_0^T u(T-t) k(t-\tau) v(\tau) d\tau dt \equiv \int_0^T v(\tau) \left[\int_0^T k(t-\tau) u(T-t) dt \right] d\tau .$$

With the substitution $T-\tau \rightarrow \sigma$ and $T-t \rightarrow \nu$ the last integral becomes

$$(4.21) \quad \int_x^0 v(T-\sigma) \left[\int_x^0 k(\sigma-\nu) u(\nu) (-d\nu) \right] (-d\sigma) \equiv \int_0^T v(T-\sigma) \left[\int_0^T k(\sigma-\nu) u(\nu) d\nu \right] d\sigma .$$

With another change (inessential) $\sigma \rightarrow t$, $\nu \rightarrow \tau$ the integral becomes

$$(4.22) \quad \int_0^T v(T-t) \left[\int_0^T k(t-\tau) u(\tau) d\tau \right] dt = \langle \mathcal{F}u, v \rangle_c$$

that shows the symmetry of the Fredholm operator with convolution kernel with respect to the convolution bilinear form.

Now a Volterra integral operator with convolution kernel can be reduced to a Fredholm integral operator with convolution kernel observing that

$$(4.23) \quad \int_0^t f(\tau) d\tau \equiv \int_0^T H(t-\tau) f(\tau) d\tau \quad (0 \leq t \leq T)$$

being $H(\xi)$ the Heaviside function

$$(4.24) \quad H(\xi) = \begin{cases} 1 & \text{for } \xi \geq 0 \\ 0 & \text{for } \xi < 0 . \end{cases}$$

Then

$$(4.25) \quad \int_0^t k(t-\tau) f(\tau) d\tau \equiv \int_0^T H(t-\tau) k(t-\tau) f(\tau) d\tau$$

and the product $H(t-\tau) k(t-\tau)$ is a new kernel of convolution kind. Then

$$(4.26) \quad \int_0^t k(t-\tau) f(\tau) d\tau = \int_0^T \bar{k}(t-\tau) f(\tau) d\tau .$$

With that reduction of the Volterra to Fredholm integral operator, our previous

result implies that also the *Volterra integral operator with convolution kernel is symmetric with the convolution bilinear form.*

This permits to give variational formulation to linear integral and integro-differential equation with convolution kernel. This fact is of interest in connection with hereditary phenomena [12] [26] [31].

4.4. Applications.

PROBLEM 1. – Give a variational formulation to the initial value problem

$$(4.27) \quad \begin{cases} \frac{d}{dt} u(t) = f(t) \\ 0 \leq t \leq T \\ u(0) = 0 \\ u(t) \in C^1[0, T]. \end{cases}$$

Since the operator is linear we can use the reduced formula (2.23) that becomes

$$(4.28) \quad F[u] = \int_{\lambda=0}^{\lambda=1} \int_0^T \left\{ \left[u(T-t) \frac{d}{dt} \lambda u(t) - u(T-t) f(t) \right] dt \right\} d\lambda \equiv \\ \equiv \frac{1}{2} \int_0^T u(T-t) u'(t) dt - \int_0^T u(T-t) f(t) dt.$$

The function $u(t)$ that solves the initial value problem (4.27) makes stationary the functional $F[u]$ given by (4.28). As a check we calculate the Euler-Lagrange equation with the « δ » process

$$(4.29) \quad \delta F[u] = \frac{1}{2} \int_0^T [\delta u(T-t) u'(t) dt + u(T-t) \delta u'(t)] dt - \int_0^T \delta u(T-t) f(t) dt \equiv \\ \equiv \frac{1}{2} \int_0^T \left[\delta u(T-t) u'(t) + \delta \frac{du(T-t)}{d(T-t)} u(t) \right] dt - \int_0^T \delta u(T-t) f(t) dt \equiv \\ \equiv \frac{1}{2} \int_0^T 2 \delta u(T-t) u'(t) dt + \frac{1}{2} [\delta u(T-t) u(t)]_0^T - \int_0^T \delta u(T-t) f(t) dt \equiv \\ \equiv \int_0^T \delta u(T-t) [u'(t) - f(t)] dt + \frac{1}{2} [\delta u(0) u(T) - \delta u(T) u(0)].$$

Since $\delta u(0) = 0$ and $\delta u(T-t)$ is arbitrary, we obtain both equations (4.27).

PROBLEM 2. - Give variational formulation to the initial value problem

$$(4.30) \quad \begin{cases} \frac{d^2}{dt^2} u(t) = f(t) \\ u(0) = 0 & 0 \leq t \leq T \\ u'(0) = 0 & u(t) \in C^2[0, 1]. \end{cases}$$

Applying the formula (2.23) one finds the functional

$$(4.31) \quad G[u] = \frac{1}{2} \int_0^T \frac{du(T-t)}{d(T-t)} \frac{du(t)}{dt} dt - \int_0^T u(T-t) f(t) dt.$$

We now show that the function $u_0(t)$ that solves the initial value problem (4.30) makes stationary the functional $G[u]$ with respect to all functions $u(t)$ that satisfy the initial condition $u(0) = 0$ and viceversa. The second initial condition $u'(0) = 0$ becomes a natural condition for the functional.

The first variation of the functional $G[u]$ is

$$(4.32) \quad \delta G[u] = \int_0^T \delta u(T-t) [u''(t) - f(t)] dt - \delta u(0) \left. \frac{du(t)}{dt} \right|_T + \delta u(T) \left. \frac{du(t)}{dt} \right|_0$$

If we require that this variation vanishes for every $u(t)$ that satisfies the condition $u(0) = 0$ we obtain the differential equation (4.30) and the initial condition $u'(0) = 0$.

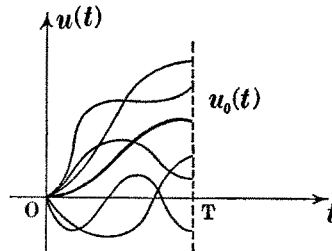


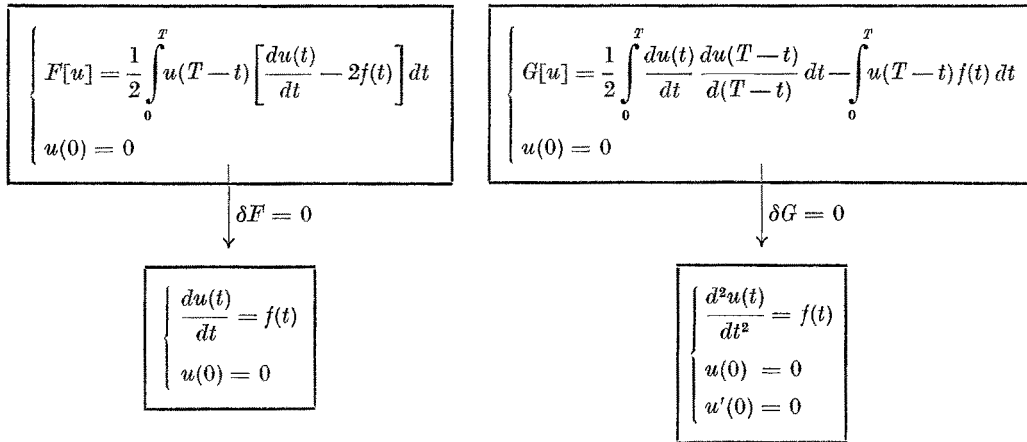
Fig. 7.

REMARK V. - If we use the usual bilinear form (4.9) we find that the operator is not symmetric while it is formally symmetric. This implies that the symmetry can be obtained changing the domain i.e. the boundary conditions: the initial condi-

tion $u'(0) = 0$ must be ignored and an artificial boundary condition $u(T) = 0$ must be added. In this way we obtain a symmetric operator. Such a variational formulation is of little use because it is limited to the fact that the Euler-Lagrange equation of such functional is the differential equation in (4.30). Since the domain

TABLE I. - *Variational formulation of the simplest initial value problems*

bilinear form $\langle u, v \rangle_c \stackrel{\text{def}}{=} \int_0^x u(T-t)v(t) dt \equiv \int_0^x u(t)v(T-t) dt$



of the functional is different from that of the ordinary problem the function $u^*(t)$ that makes the functional stationary is not, in general, a solution of the initial value problem because it does not satisfy a priori the forgotten initial condition $u'(0) = 0$.

PROBLEM 3. - Give a variational formulation to the initial value problem

$$(4.33) \quad \begin{cases} \frac{d}{dt} u(t) = 0 \\ 0 \leq t \leq T \quad u(t) \in C^1[0, T] \\ u(0) = a \quad (a \neq 0) . \end{cases}$$

On account of inhomogeneous boundary condition the operator is not linear (linearity of an operator requires the linearity of the domain). The Gateaux derivative of this operator is d/dt applied to $\varphi(t)$ such that $\varphi(0) = 0$. Then the Gateaux derivative is a linear operator that satisfies the symmetry requirement with the convolution bilinear form.

Since the domain does not contain the null element we must use relation (2.22). Putting $u_0(t) = a$ we obtain

$$(4.34) \quad H[u] = F[a] + \int_{\lambda=1}^{\lambda=0} \left\{ \int_0^T [u(T-t) - a] \frac{d}{dt} [a + \lambda(u - a)] dt \right\} d\lambda$$

that after some simplifications becomes

$$(4.35) \quad H[u] = \frac{1}{2} \int_0^T u(T-t) u'(t) dt - \frac{1}{2} au(T)$$

where we have omitted the inessential additive constant $F[a] + \frac{1}{2}a^2$. The stationarity of this functional with respect to arbitrary variations of $u(t)$ that vanish for $t = 0$ (i.e. $\delta u(0) = 0$) leads to the Euler-Lagrange equation and to the initial condition (4.33). In fact

$$(4.36) \quad \delta H[u] = \int_0^T \delta u(T-t) u'(t) dt + [\frac{1}{2}u(0) \delta u(t) - \frac{1}{2}u(T) \delta u(0)] - \frac{1}{2}a \delta u(T)$$

from which the statement follows.

PROBLEM 4. - Give a variational formulation to the linear initial value problem

$$(4.37) \quad \begin{cases} \frac{d}{dt} u_h(t) = \sum_1^n a_{hk} u_k(t) + f_h(t) \\ u_h(0) = b_h, & a_{hk} = a_{kh} \\ 0 \leq t \leq T, & u(t) \in C^1[0, T] \end{cases}$$

where a_{hk} is time-independent.

Because of the relation $a_{hk} = a_{kh}$ and of the initial conditions we can use the bilinear form

$$(4.38) \quad \langle u, v \rangle_c = \int_0^T \sum_1^n u_h(T-t) v_h(t) dt.$$

The operator

$$(4.39) \quad L = \frac{d}{dt} u_h(t) - \sum_1^n a_{hk} u_k(t)$$

is symmetric with that bilinear form and then a variational formulation can be given. The functional can be found with formula (2.22): one obtains

$$(4.40) \quad F[u] = \frac{1}{2} \int_0^T \sum_1^n u_n(T-t) \left[u_n(t) - \sum_1^n a_{nk} u_k(t) \right] dt - \\ - \int_0^T \sum_1^n u_n(T-t) f_n(t) dt - \frac{1}{2} \sum_1^n b_n u_n(T).$$

This variational formulation can be useful in linear network theory and in the theory of small vibrations around a configuration of stable equilibrium.

PROBLEM 5. — Give a variational formulation to the integral equation

$$(4.41) \quad u(t) + \int_0^t k(t-\tau) u(\tau) d\tau = f(t)$$

with convolution kernel. The operator is linear and then we can use formula (2.23). Then the functional becomes

$$(4.42) \quad F[u] = F[\theta] + \int_{\lambda=0}^{\lambda=1} \int_0^T u(T-t) \left[\lambda u(t) + \int_0^t k(t-\tau) \lambda u(\tau) d\tau \right] dt d\lambda - \int_0^T u(T-t) f(t) dt, \\ F[u] \equiv F[\theta] + \frac{1}{2} \int_0^T u(T-t) u(t) dt + \frac{1}{2} \int_0^T \int_0^T u(T-t) k(T-\tau) u(\tau) d\tau dt - \int_0^T u(T-t) f(t) dt.$$

This is the functional whose Euler-Lagrange equations is the integral equation (4.41).

5. — Comparison with other methods.

Let us compare the three methods examined at § 3 with the method exposed in this paper.

We examine the three methods in the reverse order.

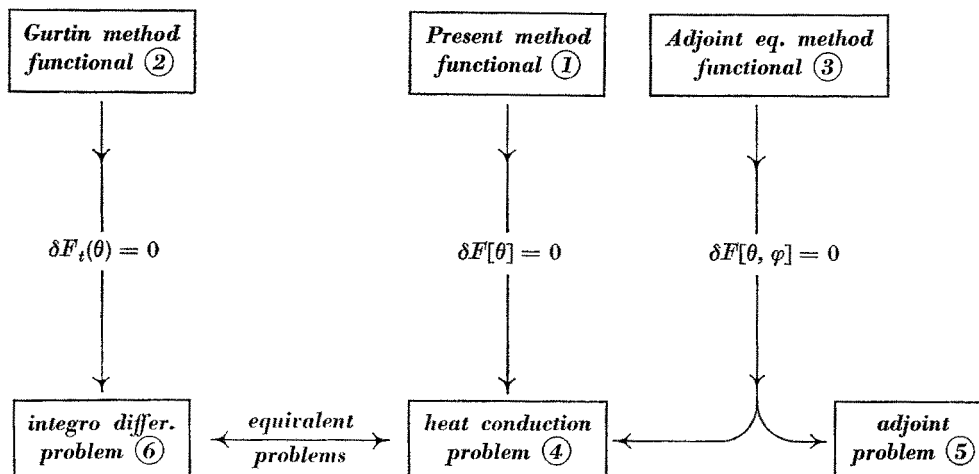
a) *Gurtin method of convolution.*

The transformation of the initial value problem into an equivalent boundary value problem governed by an integro-differential equation, that is characteristic of Gurtin method, become superfluous. We can deduce directly the initial value problem in differential form from the stationarity of a functional (see table 3).

TABLE 2. - Comparison among different methods (see next table).

①	$F[\theta] = \frac{1}{2} \int_V \int_0^T \int \left\{ \nabla \theta(\mathbf{x}, T-t) \cdot \nabla \theta(\mathbf{x}, t) + k\theta(\mathbf{x}, T-t) \frac{\partial \theta(\mathbf{x}, t)}{\partial t} \right\} dt dV - \frac{1}{2} \int_V \int k\alpha(\mathbf{x}) \theta(\mathbf{x}, t) dV$ $\theta(\mathbf{x}, 0) = \alpha(\mathbf{x}), \quad \theta(\mathbf{x}, t) _S = \beta(\mathbf{x}_S, t)$	
②	$F_t(\theta) = \frac{1}{2} \int_V \int_0^t \int \left\{ \nabla \theta(\mathbf{x}, t-\tau) \cdot \int_0^\tau \nabla \theta(\mathbf{x}, \sigma) d\sigma + k\theta(\mathbf{x}, t-\tau) \right\} d\tau dV - 2 \int_V \int_0^t \int k\alpha(\mathbf{x}) \theta(\mathbf{x}, t-\tau) d\tau dV$ $\theta(\mathbf{x}, t) _S = \beta(\mathbf{x}_S, t)$	
③	$F[\theta, \varphi] = \int_V \int_0^T \int \left\{ \nabla \varphi(\mathbf{x}, t) \cdot \nabla \theta(\mathbf{x}, t) - k\theta(\mathbf{x}, t) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} \right\} dt dV - \int_V \int k\alpha(\mathbf{x}) \varphi(\mathbf{x}, 0) dV$ $\theta(\mathbf{x}, t) _S = \beta(\mathbf{x}_S, t), \quad \varphi(\mathbf{x}, t) _S = b(\mathbf{x}_S, t), \quad \varphi(\mathbf{x}, T) = a(\mathbf{x})$	
④ Transient heat conduction problem	⑤ Adjoint problem	⑥ Integrodifferential problem equivalent to the problem ④
$\nabla^2 \theta(\mathbf{x}, t) - k \frac{\partial \theta(\mathbf{x}, t)}{\partial t} = 0$ $\theta(\mathbf{x}, 0) = \alpha(\mathbf{x})$ $\theta(\mathbf{x}, t) _S = \beta(\mathbf{x}_S, t)$	$\nabla^2 \varphi(\mathbf{x}, t) + k \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} = 0$ $\varphi(\mathbf{x}, T) = a(\mathbf{x})$ $\varphi(\mathbf{x}, t) _S = b(\mathbf{x}_S, t)$	$\int_0^t \nabla^2 \theta(\mathbf{x}, \tau) d\tau - k\theta(\mathbf{x}, t) = k\alpha(\mathbf{x})$ $\theta(\mathbf{x}, t) _S = \beta(\mathbf{x}_S, t)$

TABLE 3. - Comparison among different methods (see preceding table).



The Gurtin method is substantially simplified. In particular if we search for a variational formulation for the integro-differential equations associated with the initial value problem, using the convolution bilinear form $\langle u, v \rangle_c$ we obtain exactly Gurtin's functionals. In this sense *Gurtin method becomes a particular case of that presented here.*

b) *The adjoint equation method.*

We now show that this method is automatically absorbed in the method presented here.

Since the heat conduction problem must be valid for every $0 \leq t \leq T$, is valid also at the instant $\tau = T - t$

$$(5.1) \quad \begin{cases} \nabla^2 \theta(\mathbf{x}, T-t) - k \frac{\partial \theta(\mathbf{x}, T-t)}{\partial (T-t)} = f(\mathbf{x}, T-t) \\ \theta(\mathbf{x}, T-t)|_{\tau=t=0} = 0 \\ \theta(\mathbf{x}, T-t)|_s = 0 \end{cases}$$

where \mathbf{x} denotes the set of three variables x, y, z . Now if we put

$$(5.2) \quad \varphi(\mathbf{x}, t) \stackrel{\text{def}}{=} \theta(\mathbf{x}, T-t) \quad g(\mathbf{x}, t) \stackrel{\text{def}}{=} f(\mathbf{x}, T-t)$$

the heat conduction problem can be written

$$(5.3) \quad \begin{cases} \nabla^2 \varphi(\mathbf{x}, t) + k \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} = g(\mathbf{x}, t) \\ \varphi(\mathbf{x}, t)|_{t=T} = 0 \\ \varphi(\mathbf{x}, t)|_s = 0, \end{cases}$$

that is exactly the adjoint problem. Then the function φ whose physical meaning does not appear to be found up to now, is *the temperature θ of the same physical system evaluated at the instant $T-t$.*

Now if we ignore this identification and multiply by $\delta\varphi(\mathbf{x}, t)$ the heat equation (3.5) and by $\delta\theta(\mathbf{x}, t)$ the adjoint eq. (5.3) and integrate on $[0, T]$ and on V we obtain

$$(5.4) \quad \iiint_V \int_0^T \left\{ \delta\varphi(\mathbf{x}, t) \left[\nabla^2 \theta(\mathbf{x}, t) - k \frac{\partial \theta(\mathbf{x}, t)}{\partial t} - f(\mathbf{x}, t) \right] + \right. \\ \left. + \delta\theta(\mathbf{x}, t) \left[\nabla^2 \varphi(\mathbf{x}, T-t) + k \frac{\partial \varphi(\mathbf{x}, T-t)}{\partial t} - g(\mathbf{x}, T-t) \right] \right\} dt dV = 0.$$

After an integration by parts this functional reduces to

$$\begin{aligned}
 (5.5) \quad & - \iiint_V \int_0^T \left\{ [\nabla \delta\varphi(\mathbf{x}, t) \cdot \nabla \theta(\mathbf{x}, t) + \nabla \varphi(\mathbf{x}, t) \cdot \nabla \delta\theta(\mathbf{x}, t)] + \right. \\
 & + k \left[\delta\varphi(\mathbf{x}, t) \frac{\partial \theta(\mathbf{x}, t)}{\partial t} + \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} \delta\theta(\mathbf{x}, t) \right] + [\delta\varphi(\mathbf{x}, t) f(\mathbf{x}, t) + \delta\theta(\mathbf{x}, t) g(\mathbf{x}, t)] \left. \right\} dt dV + \\
 & + \iiint_V [\delta\varphi(\mathbf{x}, t) \mathbf{n} \cdot \nabla \theta(\mathbf{x}, t) + \delta\theta(\mathbf{x}, t) \mathbf{n} \cdot \nabla \varphi(\mathbf{x}, t)]_0^T dV = 0.
 \end{aligned}$$

If we evidentiate the « δ » operator we obtain

$$\begin{aligned}
 (5.6) \quad & - \delta \iiint_V \int_0^T \left[\nabla \varphi(\mathbf{x}, t) \cdot \nabla \theta(\mathbf{x}, t) + k\varphi(\mathbf{x}, t) \frac{\partial \theta(\mathbf{x}, t)}{\partial t} + k\theta(\mathbf{x}, t) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \right. \\
 & + \varphi(\mathbf{x}, t) f(\mathbf{x}, t) + \theta(\mathbf{x}, t) g(\mathbf{x}, t) \left. \right] dt + \iiint_V \left[\delta\varphi(\mathbf{x}, t) \frac{\partial \theta(\mathbf{x}, t)}{\partial n} + \partial\theta(\mathbf{x}, t) \frac{\partial \varphi(\mathbf{x}, t)}{\partial n} \right]_0^T dV = 0.
 \end{aligned}$$

This is the usual functional for the pair of problems (5.5) and (5.3). Now if we do the identification (5.2) and use the commutative property of the convolution, equation (5.5) becomes

$$(5.7) \quad \delta \iiint_V \int_0^T -2 \left[\nabla \theta(\mathbf{x}, T-t) \cdot \nabla \theta(\mathbf{x}, t) + k\theta(\mathbf{x}, T-t) \frac{\partial \theta(\mathbf{x}, t)}{\partial t} + \theta(\mathbf{x}, T-t) f(\mathbf{x}, t) \right] dt = 0$$

on account of

$$(5.8) \quad \iiint_V \left[\delta\theta(\mathbf{x}, T-t) \frac{\partial \theta(\mathbf{x}, t)}{\partial n} + \partial\theta(\mathbf{x}, t) \frac{\partial \theta(\mathbf{x}, T-t)}{\partial n} \right]_0^T dV \equiv 0.$$

The functional in (5.7) is exactly those that we obtain using the convolution bilinear form: its Euler-Lagrange equation is the heat conduction equation. Then the adding of the adjoint equation becomes superfluous once we have recognized that the adjoint function is no more than the same function θ evaluated at the time $T-t$. With this identification the functional obtained with that method is exactly that obtained applying directly the convolution bilinear form to the heat equation.

Then the method of adding the adjoint equation inaugurated by Morse and Feshbach has no more reason to be used.

a) The method for formally selfadjoint operators.

This method is not completely absorbed, at present, in the method presented here. The inclusion is valid for linear equations with constant coefficients.

In spite of this, such class of equations is large enough for the equations of physics. So wave equations, Maxwell equations, Schrodinger, Klein-Gordon, Dirac equations are of this kind. The nonlinear case remains to be explored. In particular we have not succeeded in doing a variational formulation that be an alternative to that of Hamilton (that deals with nonlinear differential equations) that eliminate the unpleasant features exposed in § 3.

ACKNOWLEDGMENTS. – The author is grateful to professors LIVIU SOLOMON and LUIGI AMERIO for their critical reading of the manuscript.

REFERENCES

- [1] M. M. VAINBERG, *Variational Methods in the Study of Nonlinear Operators*, Holden Day, 1964.
- [2] E. TONTI, *Variational formulation of nonlinear differential equations*, Bull. Acad. Roy. Belgique, 5^e serie, tome IV (1969), pp. 137-165 (first part), pp. 262-278 (second part).
- [3] E. TONTI, *On the Inverse Problem of the Calculus of Variations*, Academic Press (in press).
- [4] P. M. MORSE - H. FESHBACH, *Methods of Theoretical Physics*, vol. I, McGraw-Hill, 1953.
- [5] C. LANZOS, *Linear Differential Operators*, Van Nostrand, 1961.
- [6] M. BECKER, *The Principles and Applications of Variational Methods*, M.I.T. Press, 1964.
- [7] J. LEWINS, *Importance. The Adjoint Function*, Pergamon Press, 1965.
- [8] R. A. NICHOLS - S. G. BANKOFF, *Adjoint variational principles for convective diffusion*, Int. J. Heat Mass Transfer, **8**, (1963), pp. 329-335.
- [9] I. C. STATTEY, *A widely applicable type of variational integral*, Chem. Engr. Sci., **19** (1964), p. 801.
- [10] M. E. GURTIN, *Variational principles for linear initial value problems*, Quart. Appl. Math., **22** (1964), pp. 252-256.
- [11] M. E. GURTIN, *Variational principles for linear elastodynamics*, Arch. Rat. Mech., **34** (1964), p. 16.
- [12] M. E. GURTIN, *Variational principles in the linear theory of viscoelasticity*, Arch. Rat. Mech. Anal., **13** (3) (1963), pp. 179-191.
- [13] L. N. TAO, *On variational principles for electromagnetic theory*, Jour. Math. Phys., **7** (1966), p. 526.
- [14] I. HLAVACEK, *Variational principles for parabolic equations*, Aplikace Matematiky, Ceskoslovenska Akademie Ved, **14** (1969).
- [15] R. S. SANDHU - K. S. PISTER, *A variational principle for linear, coupled field problems in continuum mechanics*, Int. J. Engin. Sci., **8** (1970), pp. 989-999.
- [16] R. S. SANDHU - K. S. PISTER, *Variational principles for boundary value and initial-boundary value problems in continuum mechanics*, Int. J. Solid. Structures, **7** (1971), pp. 639-654.
- [17] N. BOURBAKI, *Eléments de mathématique, Livre V, Espaces vectoriels topologiques*, Hermann, Paris, 1966.
- [18] J. J. MOREAU, *Fonctionnelles convexes*, Séminaire sur les équations aux dérivées partielles, Collège de France (1966-67) (policopy).
- [19] J. J. MOREAU, *Convexity and duality*, in Functional analysis and optimization, Academic Press, 1966.

-
- [20] N. I. AKIEZER - I. M. GLAZMAN, *Theory of linear operators in Hilbert space*, vol. I, Ungar Publ. Co., 1961.
- [21] S. G. MIKHLIN, *Variational Methods in Mathematical Physics*, Pergamon Press, 1964.
- [22] I. HLAVACEK, *On the existence and uniqueness of solution of the Cauchy problem for linear integro-differential equations with operator coefficients*, *Apl. Matematiky*, **16** (1971), p. 64.
- [23] T. R. GOODMAN, *The adjoint heat-conduction problem for solids*, Proc. fourth U.S. Nat. Congr. Appl. Mech. 1962, vol. II, p. 1257.
- [24] R. S. SCHECHTER, *The Variational Method in Engineering*, McGraw-Hill, 1967.
- [25] H. L. DRYDEN - F. P. MURNAGHAN - H. BATEMAN, *Hydrodynamics*, Dover, 1956, p. 167.
- [26] B. D. COLEMAN, *Thermodynamics of material with memory*, *Arch. Rat. Mech. Anal.* **17** (1964), p. 1.
- [27] L. S. PONTRYAGIN, *Hermitean operators in a space with indefinite metric*, *Izv. Akad. Nauk. SSSR, Ser. Mat.*, **8**, no. 6 (1944), pp. 243-280.
- [28] R. NEVANLINNA, *Ann. Ac. Sci. Fenn.*, **1** (1953), pp. 108, 115; (1954), pp. 163-222.
- [29] L. K. PANDIT, *Linear vector spaces with indefinite metric*, *Nuovo Cimento, serie X*, **11** (1958), p. 157.
- [30] N. DUNFORD - J. T. SCHWARTZ, *Linear Operators*, Part. II, Interscience, 1963.
- [31] S. LEVONI, *Variational principles in the electrostatics of hereditary media*, *Meccanica*, **6** (1971), p. 3.
-