The Australian Journal of Mathematical Analysis and Applications

AJMAA

Volume 15, Issue 1, Article 2, pp. 1-12, 2018

INTRODUCING THE DORFMANIAN: A POWERFUL TOOL FOR THE CALCULUS OF VARIATIONS

OLIVIER DE LA GRANDVILLE

Received 8 June, 2017; accepted 5 December, 2017; published 29 January, 2018.

DEPARTMENT OF MANAGEMENT SCIENCE AND ENGINEERING, STANFORD UNIVERSITY, 475 VIA ORTEGA, STANFORD CA 94305-4121, USA
odelagrandville@gmail.com

ABSTRACT. We show how a modified Hamiltonian proposed by Robert Dorfman [1] to give intuitive sense to the Pontryagin maximum principle can be extended to easily obtain all high-order equations of the calculus of variations. This new concept is particularly efficient to determine the differential equations leading to the extremals of functionals defined by n-uple integrals, while a traditional approach would require – in some cases repeatedly – an extension of Green's theorem to n-space.
Our paper is dedicated to the memory of Robert Dorfman (1916 - 2002).

Key words and phrases: Calculus of variations; Optimal control; Modified Hamiltonian.

2010 *Mathematics Subject Classification.* 49K05, 49K15, 49K20.

ISSN (electronic): 1449-5910

1. INTRODUCTION

The calculus of variations is arguably one of the most beautiful areas of mathematics. Indeed, the problems it addressed at first seemed beyond the reach of the standard differential calculus available; then remarkably simple methods were progressively devised, illuminating areas where the best minds had only been able to grope their way. Remember the challenge offered by Johann Bernoulli, in June 1696, whose solution had eluded even Galileo: find the path $y(x)$ between two points A and B in a vertical plane such that a bead slide from A to B in minimum time. His brother Jakob, as well as Newton, Leibniz, Tschirnhaus and the Marquis de L'Hospital were able to determine that the solution, called the brachistochrone, was an arc of cycloid, but only by resorting to a series of very clever geometric and physical considerations. None of them, however, were able to find an analytic solution to the problem of extremizing an integral such as

$$I\left[y\left(x\right)\right]=\int_{a}^{b}F\left(x,y,y'\right)dx, \tag{1.1}$$

a so-called functional, i.e. a relationship between a function $y(x)$ and a number I. (The time T taken by the bead, starting at point A(x_0,y_0) to reach point B(x_1,y_1), equal to $T\left[y\left(x\right)\right]=\frac{1}{\sqrt{2g}}\int_{x_0}^{x_1}\sqrt{\frac{1+y'^2}{y_0-y}}dx$, where g is the gravitational constant, is an example of the functional (1.1), albeit the integrand does not depend explicitely on x). Mathematicians would have to wait nearly half a century for the genius of Euler to come up, in 1744, with the second-order differential equation

$$\frac{\partial F}{\partial y}(x,y,y')-\frac{d}{dx}\frac{\partial F}{\partial y'}(x,y,y')=$$

$$\frac{\partial F}{\partial y}(x,y,y')-\frac{\partial^2 F}{\partial y'\partial x}(x,y,y')-\frac{\partial^2 F}{\partial y'\partial y}(x,y,y')y'-\frac{\partial^2 F}{\partial y'^2}(x,y,y')y''=0 \tag{1.2}$$

that would constitute a first-order condition for a functional such as (1.1) to be maximized or minimized. But even Euler was not able to derive his equation from an analytic argument, and had to rely on geometrical reasoning. As Herman Goldstine relates in his splendid *History of the Calculus of Variations* [3], Euler himself tells us how he had hopelessly struggled to find such a derivation and that all the merit of achieving this feat belonged to the 19 year-old Giuseppe Ludovico de La Grange-Tournier[1]. Here are Euler's own words (quoted by Goldstine, [3], p. 110): "Even though the author of this [Euler] had meditated a long time and had revealed to friends his desire, yet the glory of first discovery was reserved to the very penetrating geometer of Turin La Grange who, having used analysis alone, has clearly attained the very same solution which the author had deduced from geometrical considerations"[2].

[1] One of the best kept secrets in mathematics is that Lagrange was Italian.

[2] Perhaps Euler had been too harsh on himself, for two reasons: first his geometric reasoning was vindicated two centuries later in the "direct methods" of the calculus of variations (see Gelfand and Fomin [2]); secondly, his approach led him to the extremal of functionals depending on *m*-order derivatives such has

$$I\left[y(x)\right]=\int_{a}^{b}F\left(x,y,y',...,y^{(m)}\right)dx,$$

given by the solution of the $2m$ order differential equation

$$\frac{\partial F}{\partial y}\left(x,y,y',...,y^{(m)}\right)+\sum_{j=1}^{m}\left(-1\right)^{j}\frac{d^j}{dx^j}\frac{\partial F}{\partial y^{(j)}}\left(x,y,y',...,y^{(m)}\right)=0,$$

Both Euler and Lagrange then dealt with more complex problems, for instance those involving constraints of varying degrees of complexity. But the work was far from finished, particularly with regard to the (admittedly thorny) sufficient conditions for the optimisation of functionals. Those would have to wait for the 19th and even the 20th century and the successive contributions of Legendre, Jacobi, Weierstrass, Hilbert and Carathéodory to be sorted out.

We need to come back for a moment to the initial problem of optimizing (1.1). We do this for two reasons. The first is to highlight the wonderful intuition of young Giuseppe; the second is to focus on *two crucial steps* that will become essential in the new discipline; indeed, it turns out that when the functional becomes complicated – for instance if it is required to maximize an n-uple integral – these steps, in the extended form they need to take, can become very intricate. On the other hand, one of the purposes of this paper is to show how the Dorfmanian – in its due extensions – avoids all of these difficulties.

WHEN YOUNG GIUSEPPE STEPPED IN

In the letter young Giuseppe sent to Euler on August 12, 1755 from his home in Torino, he suggested the following (we will use the notation corresponding to (1.1) and (1.2)). Suppose we want to find an extremal of (1.1) subject to $y(a) = y_a$ and $y(b) = y_b$. Assume that $y(x)$ has been found as an extremal. We can give to this extremal an increase – a variation, or a perturbation – with the following form: it will be the product of an *arbitrary C^1 fixed function* $\eta(x)$ such that $\eta(a) = \eta(b) = 0$, and a *variable real number* α. The new, resulting function being $y(x) + \alpha\eta(x)$, the value of the functional becomes

$$I\left[y(x) + \alpha\eta(x)\right] = \int_a^b F\left(x;\ y(x) + \alpha\eta(x);\ y'(x) + \alpha\eta'(x)\right) dx = J(\alpha). \tag{1.3}$$

Since $y(x)$ and $\eta(x)$ are *fixed functions*, the functional (1.1) that was dreaded for so many years just turned into a nice function of a single variable, here denoted $J(\alpha)$. The wonderful idea of young Giuseppe was to transform the problem of optimising a functional into a very simple problem of differential calculus whose solution could yield the Euler equation, as follows. If an interior solution exists, we must have $J'(\alpha)_{\alpha=0} \equiv J'(0) = 0$, i.e.

$$J'(0) = \int_a^b \left[\frac{\partial F}{\partial y}\eta(x) + \frac{\partial F}{\partial y'}\eta'(x)\right] dx = 0. \tag{1.4}$$

The second master stroke by La Grange was to render this first-order condition independent of $\eta(x)$. Integrating by parts the second term of the integrand – this is the first crucial step we alluded to – gives

$$\int_a^b \frac{\partial F}{\partial y'}\eta'(x)dx = \left[\frac{\partial F}{\partial y'}\eta(x)\right]_a^b - \int_a^b \eta(x)\frac{d}{dx}\frac{\partial F}{\partial y'}dx; \tag{1.5}$$

---

later called the Euler-Poisson equation. Furthermore, in the same way he was able to tackle functionals defined by double integrals such as

$$I\left[u(x,y)\right] = \int\int_R F\left(x,y,u,u_x,u_y\right) dxdy,$$

leading to the second order partial differential equation

$$\frac{\partial F}{\partial u}\left(x,y,u,u_x,u_y\right) - \frac{\partial}{\partial x}\frac{\partial F}{\partial u_x}\left(x,y,u,u_x,u_y\right) - \frac{\partial}{\partial y}\frac{\partial F}{\partial u_y}\left(x,y,u,u_x,u_y\right) = 0,$$

which would later have so much importance in physics – in particular it led to the Laplace and to the Schrödinger equations.

Since $\eta(a) = \eta(b) = 0$, the first-order condition given by (1.4) therefore becomes

$$J'(0) = \int_a^b \eta(x) \left[\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right] dx \equiv \int_a^b \eta(x) g(x) dx = 0, \tag{1.6}$$

where $g(x)$ denotes $\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'}$. With the welcome impetuosity of youth, La Grange concluded that for the integral to vanish, definitely $g(x)$ should be equal to zero over the whole interval $[a, b]$ since $\eta(x)$ was to be arbitrary. Euler agreed to this, but did ask Giuseppe to please prove it. That proof became known as the fundamental lemma of the calculus of variations. This is the second step we alluded to, and the final part of the analytic derivation of the Euler formula.

Lemma 1.1. *If $g(x)$ is continuous in $[a, b]$ and if $\int_a^b \eta(x) g(x) dx = 0$ for any function $\eta(x) \in C(a, b)$ such that $\eta(a) = \eta(b) = 0$, then $g(x) = 0$ for all x in $[a, b]$.*

Proof. Suppose that, at some point c in $[a, b]$, $g(c)$ is, say, positive. Since $g(x)$ is supposed to be continuous over $[a, b]$, definitely there must exist, around c and, included in $[a, b]$, an interval $[c_1, c_2]$ where $g(x)$ is positive. Suppose on the other hand that $g(c)$ vanishes on the remaining intervals. Therefore, we must have

$$\int_a^b \eta(x) g(x) dx = \int_{c_1}^{c_2} \eta(x) g(x) dx > 0, \tag{1.7}$$

which proves the lemma by contradiction. Hence we must have $g(x) \equiv \frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} = 0$ and the Euler equation (1.2) is validated as a first order condition for the optimization of functional (1.1). ∎

*

* *

As mentioned above, these steps – in extended forms – become unescapable whenever we want to deal, in the classical calculus of variations, with more complicated functionals such as integrals depending on higher derivatives or multiple integrals. In the case of *n*-uple integrals, the extension of the fundamental lemma of the calculus of variations is easy to conceive and demonstrate; but if, in the case of double integrals, it is easy to extend the concept of integration by parts in the form of Green's theorem, that is definitely not the case if we deal with n-uple integrals (see for instance Gelfand and Fomin [2], pp. 153-154, or Troutman [9], pp. 179-181). We will now show how the new concept introduced by Robert Dorfman enables, in extended forms, to eschew these hurdles and obtain the general Euler equations in a very straightforward, simple way. To do that we need to make a little detour by paying a short visit to a simple problem in optimal control theory.

In the middle of the 20th century, significant developments were brought to the calculus of variations. Originally motivated by the need to define and control optimal time paths in aero- and astro-engineering[3], the theory of optimal control finds today huge applications in countless scientific endeavours; even in its early stages, it was in the tool-kit of some economists interested in defining optimal trajectories for the economy.

[3] The birth of optimal control can be assigned to two important papers by Magnus Hestenes who was working at the time at the RAND Corporation [4, 5]. Hestenes enlarged the basic assumptions of the calculus of variations and introduced the distinction between state variables and control variables. Other prominent researchers at Rand at the same time were Richard Bellman, the father of dynamic programming, and Rufus Isaacs to whom we owe the concept of differential games. The modern presentation of optimal control theory used today under the form of the Pontryagin maximum principle is due to four eminent mathematicians from the Steklov Mathematics Institute (Moscow): Lev Pontryagin, Vladimir Boltyanskii, Revaz Gamkrelidze and Evgenii Mischenko. Their book [8] appeared in Russian in 1961; it was translated into English in 1962 and into German in 1964. A fascinating account of the birth of optimal control during the cold war is in Pesch and Plail [6, 7]

It is precisely on the basis of a very simple economic model that Robert Dorfman was able to achieve two things: first to rationalize the main result of optimal control theory, known as the Pontryagin maximum principle; and second to suggest a modified Hamiltonian that would yield the basic principle directly.

In Section 2 we recall Robert Dorfman's remarkable ideas. In Section 3 we will show how fruitful they can be.

2. DORFMAN'S DISCOVERY

Consider the very simple, following problem. Suppose a firm obtains at time t, over an infinitesimal duration dt, a flow of net income u which is a function of three variables

$$u = u(k_t, x_t, t) \tag{2.1}$$

where k_t and x_t are defined as follows:

. k_t is the capital stock of the firm at time t (this capital stock is the aggregate value of its buildings, equipment, machinery, etc); k_t is called a "state" variable, reflecting one of the characteristics of the firm at that instant;

. x_t is a "control" variable that reflects decisions taken by the firm at time t; for instance, those may be investment that will eincrease the capital stock, or an enhancement of the quality of the product.

A fundamental hypothesis is that the time path of the state variable variable k_t is governed by the first order differential equation

$$\dot{k} = f(k_t, x_t, t). \tag{2.2}$$

Suppose now that with an initial condition on k we know the optimal time path of the control variable x_t over time interval $[t_0, T]$. Then system (2.1),(2.2) is liable to generate the corresponding trajectories of capital k and net income u. If an aim of the firm is to maximize the total of its net income over that period, it has to determine the trajectory of x_t maximizing the functional

$$I = \int_{t_0}^{T} u(k_t, x_t, t)dt, \tag{2.3}$$

subject to (2.2). (Note that the dependency of function $u(k_t, x_t, t)$ on t may reflect factors as diverse as technical progress as well as the discounting of future income flows). In this very simple case, the Pontryagin principle would require to form a Hamiltonian function defined as

$$H = u(k_t, x_t, t) + \lambda(t) f(k_t, x_t, t) \tag{2.4}$$

where a new variable $\lambda(t)$, called an adjoint, or costate variable appears as a multiplier of $\dot{k} = f(k_t, x_t, t)$; its signification will soon appear. In its very simplified form (because we do not have here possible complexities of the general principle such as inequality constraints or the non-differentiability of (2.4) with respect to the control x_t), a first-order condition for the trajectory of x_t to be optimal is that $k_t, x_t,$ and $\lambda(t)$ solve the system

$$\frac{\partial H}{\partial x} = \frac{\partial}{\partial x} u(k_t, x_t, t) + \lambda(t) \frac{\partial}{\partial x} f(k_t, x_t, t) = 0 \tag{2.5}$$

$$\frac{\partial H}{\partial k} = \frac{\partial}{\partial k} u(k_t, x_t, t) + \lambda(t) \frac{\partial}{\partial k} f(k_t, x_t, t) = -\dot{\lambda} \tag{2.6}$$

$$\frac{\partial H}{\partial \lambda} = \dot{k} \tag{2.7}$$

If (2.7) is just a consequence of the properties of (2.2) an (2.4), the signification of equations (2.5) and (2.6) is far from obvious. Dorfman first provided an intuitive justification of these equations by reasoning as follows. Consider the problem of maximizing integral I over any interval $[t, T]$; I can be written in two parts:

$$I = \int_t^T u(k_\tau, x_\tau, \tau)d\tau = u(k_t, x_t, t)\Delta t + \int_{t+\Delta t}^T u(k_\tau, x_\tau, \tau)d\tau \tag{2.8}$$

where Δt is an extremely small time span. Suppose now that the optimal trajectory of $x,$ denoted x^*, has been found from time $t + \Delta t$ to time T. Furthermore, we observe that at any point of time $t,$ the value of x_t has an impact on two magnitudes: first, on the integrand $u(k_t, x_t, t);$ second, on the value of $k_{t+\Delta t}$, through the differential equation $\dot{k} = f(k_t, x_t, t);$ we have, in linear approximation, $k_{t+\Delta t} \approx k_t + f(k_t, x_t, t)\Delta t$. From the latter impact we can conclude that the *optimal value* of the integral on the right-hand side of (2.8) depends solely on x_t – through the variable $k_{t+\Delta t}$. It can thus be denoted as

$$\int_{t+\Delta t}^T u(k_\tau, x_\tau, \tau)d\tau \equiv V^*_{t+\Delta t}\left[k_{t+\Delta t}\left(x_t\right)\right]; \tag{2.9}$$

therefore the functional I has turned into a nice function depending solely on the single variable, x_t (young Giuseppe might have approved of this); this function reads

$$I\left(x_t\right) = u(k_t, x_t, t)\Delta t + V^*_{t+\Delta t}\left[k_{t+\Delta t}\left(x_t\right)\right]. \tag{2.10}$$

Taking the derivative of $I\left(x_t\right)$ to zero gives

$$I'\left(x_t\right) = \frac{\partial}{\partial x}u(k_t, x_t, t)\Delta t + \frac{\partial V^*_{t+\Delta t}}{\partial k_{t+\Delta t}}\frac{\partial k_{t+\Delta t}}{\partial x_t} = 0. \tag{2.11}$$

The term $\partial V^*_{t+\Delta t}/\partial k_{t+\Delta t}$ has considerable signification; it is the rate of increase of the optimal value $V^*_{t+\Delta t}$ per additional unit of capital available at $t + \Delta t$; it is thus the *price* of one unit of capital available at time $t + \Delta t$; this price can be denoted $\lambda(t + \Delta t)$. On the other hand, $\partial k_{t+\Delta t}/\partial x_t$ is, in linear approximation, $\left[\partial f(k_t, x_t, t)/\partial x_t\right]\Delta t$; therefore (2.11) can be written

$$\frac{\partial}{\partial x}u(k_t, x_t, t)\Delta t + \lambda(t + \Delta t)\frac{\partial f(k_t, x_t, t)}{\partial x_t}\Delta t = 0. \tag{2.12}$$

Dividing by Δt and taking the limit when $\Delta t \to 0$ gives equation (2.5), the first equation of the Pontryagin principle.

Suppose that this first-order condition, associated to other second-order conditions, enabled us to maximize the integral $I = \int_t^T u(k_\tau, x_\tau, \tau)d\tau$, which can now be denoted $V_t^*(k_t)$. Equation (2.8) becomes the identity

$$V_t^*(k_t) = u(k_t, x_t, t)\Delta t + V^*_{t+\Delta t}(k_{t+\Delta t}); \tag{2.13}$$

differentiating (2.13) with respect to k_t and using $\partial V_t^*/\partial k_t = \lambda(t)$ gives
(2.14)

$$\lambda(t) = \frac{\partial}{\partial k_t}u(k_t, x_t, t)\Delta t + \frac{\partial V^*_{t+\Delta t}(k_{t+\Delta t})}{\partial k_{t+\Delta t}}\frac{\partial k_{t+\Delta t}}{\partial k_t} = \frac{\partial}{\partial k_t}u(k_t, x_t, t)\Delta t + \lambda(t + \Delta t)\frac{\partial k_{t+\Delta t}}{\partial k_t}.$$

Using the linear approximations $\lambda(t + \Delta t) \approx \lambda(t) + \dot{\lambda}(t)\Delta t$ and $k_{t+\Delta t} \approx k_t + f(k_t, x_t, t)\Delta t$, (2.14) becomes

$$\lambda(t) = \frac{\partial}{\partial k_t}u(k_t, x_t, t)\Delta t + \left[\lambda(t) + \dot{\lambda}(t)\Delta t\right]\left[1 + \frac{\partial}{\partial k_t}f(k_t, x_t, t)\Delta t\right] =$$

$$\frac{\partial}{\partial k_t}u(k_t,x_t,t)\Delta t+\lambda(t)+\dot{\lambda}(t)\Delta t+\lambda(t)\frac{\partial}{\partial k_t}f(k_t,x_t,t)\Delta t+\dot{\lambda}(t)\frac{\partial}{\partial k_t}f(k_t,x_t,t)(\Delta t)^2; \tag{2.15}$$

cancelling $\lambda(t)$ leads in the limit to

$$\frac{\partial}{\partial k_t}u(k_t,x_t,t)+\lambda(t)\frac{\partial}{\partial k_t}f(k_t,x_t,t)\Delta t=-\dot{\lambda}(t), \tag{2.16}$$

the second equation of the Pontryagin principle (2.6).

Dorfman already achieved two things: he provided an intuitive justification of those equations that may look somewhat abtruse to many – particularly equation (2.6); in addition, he gave a clear signification to the adjoint variable $\lambda(t)$ as the *value* of one additional unit of the state variable at any time t.

*
* *

It is at this point that Dorfman had a remarkable insight: he saw that once we had secured the signification of the adjoint variable $\lambda(t)$ there was an even simpler route to derive these equations. The key was to introduce what he called a "modified Hamiltonian". To pay tribute to Robert Dorfman as well as to honor his memory, we propose to call this new concept a *Dorfmanian.* This new concept, denoted H^* by Dorfman, and D by us, is defined as follows:

$$D=H+\dot{\lambda}(t)k_t=u\left(k_t,x_t,t\right)+\lambda(t)\dot{k}_t+\dot{\lambda}(t)k_t=u\left(k_t,x_t,t\right)+\frac{d}{dt}\left[\lambda(t)k_t\right]. \tag{2.17}$$

The Dorfmanian is thus the traditional Hamiltonian augmented by $\dot{\lambda}(t)k_t$, i.e. the change in the value of the state variable due to the sole change in the unit price of k. This new, augmented Hamiltonian has an immediate signification. At any point of time t, the decision taken on the control variable is exerting two effects on the value of the integral I to be maximized:

a) a direct one through the integrand $u\left(k_t,x_t,t\right)$

b) an indirect one, equal to the rate of increase in the value of the state variable $\frac{d}{dt}\left[\lambda(t)k_t\right]=\lambda(t)\dot{k}_t+\dot{\lambda}(t)k_t$.

The Dorfmanian is the exact measure of the sum of those two effects. It is then obvious that setting to zero the gradient of D with respect to x_t and k_t will lead to the system of equations whose solution will be a candidate to be an extremal of I. From a methodological point of view, we can see that x_t and k_t are now treated in a symmetrical way; although k_t is not a control variable, it nonetheless is a variable that is indirectly controled through x_t. As such it now plays the same role as x_t when the gradient of D is taken to zero.

Before turning to applications of the Dorfmanian to high-order problems of the calculus of variations we can already see how very simple the determination of the Euler equation, for example, will be. Consider the minimisation of $\int_a^b F(x,y,y')dx$ conditional to the usual boundary conditions. The Dorfmanian will be $D=F(x,y,y')+\lambda(x)y'+\lambda'(x)y$; $D_y=0$ and $D_{y'}=0$ give $F_y+\lambda'(x)=0$ and $F_{y'}+\lambda(x)=0$, from which $F_y-\frac{d}{dx}F_{y'}=0$ immediately results.

3. EXTENSIONS

We will now show how the Dorfmanian can be extended in a natural way to solve variational problems involving higher order functionals. We will consider integrals whose integrands depend on several functions; functionals depending on high-order derivatives; and functionals defined by multiple integrals.

3.1. **Functionals depending on n functions $y_1, ..., y_n$.** Suppose our functional is defined by

$$I[y_1, ..., y_n] = \int_a^b F\left(x, y_1, ..., y_n, y_1', ..., y_n'\right) dx. \tag{3.1}$$

with $2n$ boundary conditions

$$y_1(a) = y_{1a},\, y_1(b) = y_{1b}$$
$$\ldots$$
$$y_n(a) = y_{na},\, y_n(b) = y_{nb}.$$

We introduce n adjoint functions $\lambda_i(x)$, $i = 1, ..., n$, each of them pertaining to state variable y_i, and playing a role anologous to that of $\lambda(x)$ in Section 1. The Dorfmanian will become

$$D = F\left(x, y_1, ..., y_n, y_1', ..., y_n'\right) + \sum_{i=1}^{n} \frac{d}{dx}\left[\lambda_i(x)\, y_i(x)\right] =$$

$$F\left(x, y_1, ..., y_n, y_1', ..., y_n'\right) + \sum_{i=1}^{n}\left[\lambda_i'(x)\, y_i(x) + \lambda_i(x)\, y_i'(x)\right]. \tag{3.2}$$

Setting to zero the gradient of D with respect to the $2n$ variables y_i, y_i', $i = 1, ..., n$ leads to

$$\frac{\partial D}{\partial y_i} = \frac{\partial F}{\partial y_i}\left(x, y_1, ..., y_n, y_1', ..., y_n'\right) + \lambda_i'(x) = 0, \quad i = 1, ..., n$$

$$\frac{\partial D}{\partial y_i'} = \frac{\partial F}{\partial y_i'}\left(x, y_1, ..., y_n, y_1', ..., y_n'\right) + \lambda_i(x) = 0, \quad i = 1, ..., n.$$

With

$$\lambda_i'(x) = -\frac{d}{dx}\frac{\partial F}{\partial y_i'}\left(x, y_1, ..., y_n, y_1', ..., y_n'\right) = 0, \quad i = 1, ..., n$$

we immediately get the system of n Euler equations

$$\frac{\partial D}{\partial y_i} = \frac{\partial F}{\partial y_i}\left(x, y_1, ..., y_n, y_1', ..., y_n'\right) - \frac{d}{dx}\frac{\partial F}{\partial y_i'}\left(x, y_1, ..., y_n, y_1', ..., y_n'\right) = 0, \quad i = 1, ..., n. \tag{3.3}$$

3.2. **Functionals depending on *m*-order derivatives.** Let $F\left(x, y, y', ..., y^{(m)}\right)$ be a function with continuous derivatives up to the second order with respect to all its arguments. Suppose we want to determine an extremal $y(x)$ of

$$I[y(x)] = \int_a^b F\left(x, y, y', ..., y^{(m)}\right) dx.$$

with $2m$ boundary conditions

$$y(a) = y_a;\, y(b) = y_b$$
$$y'(a) = y_a';\, y'(b) = y_b'$$
$$\ldots \qquad \ldots$$
$$y^{(m-1)}(a) = y_a^{(m-1)};\, y^{(m-1)}(b) = y_b^{(m-1)}$$

applying to the admissible functions as well as to all their derivatives until their $(m-1)$th order inclusively. Classically, the variation of the functional involves $m+1$ terms of the form $F_{y^{(j)}}\eta^{(j)}(x)$, $j=0,...,m$ (with the notation $y \equiv y^{(0)}$). The second term is to be integrated once by parts, the j-th will be integrated $j-1$ times by parts, and the last term, m times. This implies, at least conceptually, $m(m+1)/2$ integrations by parts. On the other hand, using the Dorfmanian will necessitate only a few derivations as we will now see.

Consider m adjoint functions $\mu_j(x)$, $j=0,...,m-1$ analogous to those introduced earlier. The Dorfmanian for this problem will be

$$D = F\left(x, y, y', ..., y^{(m)}\right) + \sum_{j=0}^{m-1} \frac{d}{dx}\left[\mu_j(x)\, y^{(j)}(x)\right] =$$

$$F\left(x, y, y', ..., y^{(m)}\right) + \sum_{j=0}^{m-1}\left[\mu_j'(x)\, y^{(j)}(x) + \mu_j(x)\, y^{(j+1)}(x)\right]. \tag{3.4}$$

Taking to zero the gradient of D with respect to $y, y', ..., y^{(m)}$ leads to a system of m equations from which $\mu_{m-1}'(x)$ and $\mu_{m-2}'(x)$ can be identified as

$$\mu_{m-1}'(x) = -\frac{d}{dx}\frac{\partial F}{\partial y^{(m)}}\left(x, y, y', ..., y^{(m)}\right)$$

and

$$\mu_{m-2}'(x) = -\frac{d}{dx}\frac{\partial F}{\partial y^{(m-1)}}\left(x, y, y', ..., y^{(m)}\right) + \frac{d^2}{dx^2}\frac{\partial F}{\partial y^{(m)}}\left(x, y, y', ..., y^{(m)}\right)$$

Proceeding successively in this way gives $\mu_0'(x)$ as

$$\mu_0'(x) = \sum_{j=1}^{m}(-1)^j \frac{d^j}{dx^j}\frac{\partial F}{\partial y^{(j)}}\left(x, y, y', ..., y^{(m)}\right)$$

and immediately the $2m$ order Euler-Poisson equation

$$\frac{\partial F}{\partial y}\left(x, y, y', ..., y^{(m)}\right) + \sum_{j=1}^{m}(-1)^j \frac{d^j}{dx^j}\frac{\partial F}{\partial y^{(j)}}\left(x, y, y', ..., y^{(m)}\right) = 0 \tag{3.5}$$

3.3. **Extension to the extremum of an $n-$tuple integral.** The advantage of having recourse to the Dorfmanian rather than going through the traditional approach becomes even more obvious if we consider the more difficult problem of finding extremals to functionals defined by an $n-$tuple integral such as

$$I\left[z(x_1, ..., x_n)\right] = \int\int\cdots\int_R F(x_1, ..., x_n,\ z,\ z_{x_1}, ..., z_{x_n})dx_1...dx_n; \tag{3.6}$$

In this problem, as usual we suppose that the domain R and its boundary ∂R have sufficient regularity. The traditional method of deriving the Euler-Ostrogradski equation requires a generalisation of Green's theorem in n-space. Contrast this with the simplicity of writing a modified Dorfmanian as follows:

$$D = F(x_1, ..., x_n, z, z_{x_1}, ..., z_{x_n}) + \sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left[\lambda(x_1, ..., x_i, ..., x_n) z(x_1, ..., x_i, ..., x_n)\right] =$$

$$F(x_1, ..., x_n,\ z,\ z_{x_1}, ..., z_{x_n}) + z\sum_{i=1}^{n}\lambda_{x_i} + \lambda\sum_{i=1}^{n} z_{x_i}\ . \tag{3.7}$$

Setting to zero the gradient of D with respect to the $n+1$ variables $z,\ z_{x_1}, ...z_{x_n}$, we get the system of $n+1$ equations

$$\frac{\partial D}{\partial z} = \frac{\partial F}{\partial z} + \sum_{i=1}^{n}\lambda_{x_i} = 0 \tag{3.8}$$

and

$$\frac{\partial D}{\partial z_{x_i}} = \frac{\partial F}{\partial z_{x_i}} + \lambda(x_1, ..., x_i, ..., x_n) = 0,\ i = 1, ..., n. \tag{3.9}$$

The last n equations yield

$$\lambda_{x_i} = -\frac{\partial}{\partial x_i}\frac{\partial F}{\partial z_{x_i}},\ i = 1, ..., n; \tag{3.10}$$

replacing these values into (2.6) gives

$$\frac{\partial F}{\partial z} - \sum_{i=1}^{n}\frac{\partial}{\partial x_i}\frac{\partial F}{\partial z_{x_i}} = 0, \tag{3.11}$$

the general Ostrogradski (or Euler) equation, obtained in four lines, without any recourse to a generalisation of Green's theorem.

If we consider functionals represented by multiple integrals whose integrand depends on higher order derivatives, as before the Dorfmanian can be generalized in a straightforward way. To illustrate, suppose we want to find an extremal $u(x, y)$ of the following integral

$$I\left[u(x,y)\right] = \int\int_R F\left(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}\right) dxdy. \tag{3.12}$$

We now introduce three adjoint functions $\lambda(x, y)$, $\mu(x, y)$ and $\omega(x, y)$ corresponding to the variables u, u_x and u_y respectively. The relevant Dorfmanian is

$$\begin{aligned} D = {} & F\left(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}\right) + \frac{\partial}{\partial x}\left[\lambda(x,y)u(x,y)\right] + \frac{\partial}{\partial y}\left[\lambda(x,y)u(x,y)\right] + \\ & \frac{\partial}{\partial x}\left[\mu(x,y)u_x(x,y)\right] + \frac{\partial}{\partial y}\left[\mu(x,y)u_x(x,y)\right] + \frac{\partial}{\partial x}\left[\omega(x,y)u_y(x,y)\right] + \frac{\partial}{\partial y}\left[\omega(x,y)u_y(x,y)\right] \end{aligned} \tag{3.13}$$

and, in simplified notation, reads

$$\begin{aligned} D = {} & F\left(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}\right) + \\ & + (\lambda_x + \lambda_y)u + (\lambda + \mu_x + \mu_y)u_x + (\lambda + \omega_x + \omega_y)u_y + \mu u_{xx} + (\mu + \omega)u_{yx} + \omega u_{yy} \end{aligned} \tag{3.14}$$

Setting the gradient of D to zero, we get

$$\frac{\partial D}{\partial u} = F_u + \lambda_x + \lambda_y = 0 \tag{3.15}$$

$$\frac{\partial D}{\partial u_x} = F_{u_x} + \lambda + \mu_x + \mu_y = 0 \tag{3.16}$$

$$\frac{\partial D}{\partial u_y} = F_{u_y} + \lambda + \omega_x + \omega_y = 0 \tag{3.17}$$

$$\frac{\partial D}{\partial u_{xx}} = F_{u_{xx}} + \mu = 0 \tag{3.18}$$

$$\frac{\partial D}{\partial u_{xy}} = F_{u_{xy}} + \mu + \omega = 0 \tag{3.19}$$

$$\frac{\partial D}{\partial u_{yy}} = F_{u_{yy}} + \omega = 0. \tag{3.20}$$

To eliminate the adjoint functions $\lambda(x, y), \mu(x, y)$ and $\omega(x, y),$ we differentiate once equations (3.16) and (3.17) with respect to x and $y,$ respectively, and in an analogous way take the second derivatives of (3.18), (3.19) and (3.20). We get the system

$$F_u + \lambda_x + \lambda_y = 0 \tag{3.21}$$

$$\frac{\partial}{\partial x}\frac{\partial D}{\partial u_x} = \frac{\partial}{\partial x}F_{u_x} + \lambda_x + \mu_{xx} + \mu_{yx} = 0 \tag{3.22}$$

$$\frac{\partial}{\partial y}\frac{\partial D}{\partial u_y} = \frac{\partial}{\partial y}F_{u_y} + \lambda_y + \omega_{xy} + \omega_{yy} = 0 \tag{3.23}$$

$$\frac{\partial^2}{\partial x^2}\frac{\partial D}{\partial u_{xx}} = \frac{\partial^2}{\partial x^2}F_{u_{xx}} + \mu_{xx} = 0 \tag{3.24}$$

$$\frac{\partial^2}{\partial x\partial y}\frac{\partial D}{\partial u_{xy}} = \frac{\partial^2}{\partial x\partial y}F_{u_{xy}} + \mu_{xy} + \omega_{xy} = 0 \tag{3.25}$$

$$\frac{\partial^2}{\partial y^2}\frac{\partial D}{\partial u_{yy}} = \frac{\partial^2}{\partial y^2}F_{u_{yy}} + \omega_{yy} = 0 \tag{3.26}$$

Upon summing equations (3.25) – (3.26) and then (3.20), (3.21), the general Euler 4th-order partial differential equation

$$F_u - \frac{\partial}{\partial x}F_{u_x} - \frac{\partial}{\partial y}F_{u_y} + \frac{\partial^2}{\partial x^2}F_{u_{xx}} + \frac{\partial^2}{\partial x\partial y}F_{u_{xy}} + \frac{\partial^2}{\partial y^2}F_{u_{yy}} = 0 \tag{3.27}$$

can be immediately derived[4]. This equation can be reached in a classical way after quite a lengthy process only, since it involves the repeated use of Green's theorem, not to mention an extension of the fundamental lemma of the calculus of variations. By contrast, we showed that the Dorfmanian led to it in just a few lines.

[4] The simplicity – or is it the elegance? – of this general Euler equation should not hide its complexity. Indeed, it contains no less than *186 terms*, each of them depending on the 8 variables $x, y, u, u_x, u_y, u_{xx}, u_{xy}$ and u_{yy}; among those 186 terms, 9 include a 4th-order partial derivative of $u(x, y)$, and 75 exibit a 3rd-order partial derivative (for example, $\frac{d^2}{dx^2}F_{u_{xx}}$ alone contains 55 terms; 3 of them carry fourth-order partial derivatives and 23 include third-order ones.

4. CONCLUDING REMARKS

There are, in our opinion, two main reasons for using the modified Hamiltonian suggested by Robert Dorfman. The first is its immediate signification as encompassing both the direct and indirect effects of the control variable. The direct effect is the integrand in the functional to be maximized or minimized; the indirect effect is the rate of change in the value of the state variables resulting from the values taken by the control variables. It then makes a lot of sense that the sum of those effects should be maximized or minimized. The second reason is that, apart from yielding the Euler equation in just one line, the Dorfmanian, in any of its extensions, offers a remarkably simple way to solve high-order variational problems, particularly those where the functional is a multiple integral.

REFERENCES

[1] ROBERT DORFMAN, An economic interpretation of optimal control theory, *American Economic Review*, 59, (5), (1969), pp. 817-31.

[2] I. GELFAND and S. FOMIN, *Calculus of Variations*, Prentice Hall, Englewood Cliffs, NJ (1961).

[3] H. GOLDSTINE, *A History of the Calculus of Variations*, Springer, 1980.

[4] M. HESTENES, Numerical methods for obtaining solutions of fixed end point problems in the calculus of variations, Research Memorandum No. 102, RAND Corporation, Santa Monica (1949).

[5] M. HESTENES, A general problem in the calculus of variations with applications to the paths of least time, Research Memorandum, Document AD 112382, RAND Corporation, Santa Monica (1949).

[6] H. J. PESCH and M. PLAIL, The maximum principle of optimal control: a history of ingenious ideas and missed opportunities, *Control and Cybernetics*, 38, no. 4A, pp. 973-995.

[7] H. J. PESCH and M. PLAIL, The Cold War and the maximum principle of optimal control, *Document Mathematica*, Extra Volume, ISMP, 2012, pp. 331-343.

[8] L. S. PONTRYAGIN, V. G. BOLTYANSKII, R. V. GAMKRELIDZE and E. F. MISCHENKO, :if expand("The Mathematical Theory of Optimal Processes (trans. K. N. Trirogoff), Interscience, New York (1962).

[9] JOHN TROUTMAN, *Variational Calculus with Elementary Convexity*, Springer Verlag, New York, 1983.