GENERALIZING VARIATIONAL THEORY TO INCLUDE THE INDEFINITE INTEGRAL, HIGHER DERIVATIVES, AND A VARIETY OF MEANS AS COST VARIABLES

JOHN GREGORY†

Abstract. In this paper we generalize the Calculus of Variations setting from the arguments \( t, x(t), x'(t) \) to the arguments \( t, x(t), x'(t), \hat{x}(t) \), where \( \hat{x}(t) \) is any indefinite integral of \( x(t) \) or its extension. Related results are to various types of “averaging” variables and to higher derivatives.

Of interest is that this process extends the basic necessary condition, the Euler equation(s), from a second order ordinary differential equation (ODE) to a second order differential equation involving indefinite integrals. These ideas can be extended to optimal control and other constraint optimization problems whose trajectory includes \( \hat{x} \) and/or averaging operators.

Of special interest is that these results can be obtained either by using a new extension of classical arguments or by a new use of the author’s constraint optimization theory. Finally, these multi-integral problems can now be solved by efficient numerical methods, previously developed by the author, with a global a priori error estimate of \( O(h^2) \).

Key words. Generalized variational theory, new dependent variables, double indefinite integrals.

AMS subject classifications. 49J15, 49K15, 34K35, 65K10

1. Introduction. The main purpose of this paper is to extend the ideas and methods of the calculus of variations/optimal control theory to problems which include the antiderivative and related dependent variables. Our basic problem is given as

\[
\min \int_a^b f(t, x(t), x'(t), \hat{x}(t)) \, dt
\]

s.t. \( x(a) = x_a \) and \( x(b) = x_b \) where

\[ \hat{x}(t) = \int_a^t x(t) \, dt. \] (1.1)

For a multitude of reasons we also assume that \( f_{x'x'} > 0 \) (see, for example, [1], [5] or [8]).

We assume this problem has a unique solution. “Smoothness” conditions will be obvious and assumed as needed. The initial conditions \( x(a) \) and \( x(b) \) can be replaced by the usual transversality conditions if these values are unspecified.

We note that our solution for (1.1) gives a more general theory for the usual classical calculus of variations/optimal control theory. In particular, the major necessary condition, the Euler equation, is no longer an ordinary second order differential equation but, now, an equation with indefinite integrals. This often implies the equation is essentially a third or fourth order ODE.

In Sections 2 and 3 we derive our key results using two distinct methods. In part this is due to the fact that our methods are themselves new and these results are so new that we wish to double check them. An example is given at the end of Section 2 which is then revisited after Section 3.

In Section 4 we extend these ideas to optimal control theory where both the cost functional and the trajectory equation include the new variable, \( \hat{x} \). We will show that the Pontryagin Maximal Principle is a special case of our results.

*Received May 31, 2007; accepted for publication October 14, 2008.
†Department of Mathematics, Mailcode 4408, Southern Illinois University Carbondale, 1245 Lincoln Drive, Carbondale IL 62901, USA (jgregory@math.siu.edu).
In Section 5 we study the case where \( f(t, x, x', x'') \) with natural/sensible boundary conditions. The extension to higher derivatives, i.e., \( f(t, x, x', \ldots, x^{(n)}) \) where \( n > 2 \) is immediate.

In Section 6 we consider the general case where \( x_2(t) = g(t) \int_a^t x_1(s) \, ds \) and in particular, \( g(t) = \frac{1}{t-a} \) to get the “running mean” as a cost variable.

2. Constraint Optimization Methods. The purpose of this section is to solve Problem (1.1) using the previous constraint methods of the author ([5] and [7]). At the end of this section we include an example to clarify the theory. A second distinct method of solution using an extension of classical variational theory methods yielding the main results as this section, will be given in Section 3.

Thus, let

\[
\min \int_a^b f(t, x, x', \hat{x}) \, dt
\]

where \( \hat{x} = \int_a^t x(s) \, ds \), and

\[
x(a) = x_a, \ x(b) = x_b.
\]

As noted above, we assume \( f_{x'x'} > 0 \). Let \( x_1 = x \), \( x_2 = \int_a^t x_1(s) \, ds \), \( x'_3 \) be the multiplier, \( X(t) = (x_1(t), x_2(t), x_3(t))^T \), and

\[
F(t, X, X') = f(t, x_1, x_1', x_2) + x'_3(x'_2 - x_1), \quad X(a) = \begin{pmatrix} x_a \\ 0 \\ 0 \end{pmatrix}, \quad X(b) = \begin{pmatrix} x_b \\ * \\ * \end{pmatrix},
\]

(2.2)

The first 0 in \( X(a) \), \( X_2(a) \), is the constant of integration. It and the constraint \( x'_2 - x_1 = 0 \) are equivalent to the integral definition of \( x_2 \). The second 0 in \( X(a) \), \( X_3(a) \) is to “normalize” \( x_3(t) \) which otherwise is not unique, although, \( x'_3(t) \) is unique. The “\(*\)” in \( X(b) \) is because \( x_2(b) \) and \( x_3(b) \) are not specified, so the usual transversality conditions \( x'_3(b) = 0 \) and \((x'_2 - x_1)(b) = 0 \) hold. From [5] or [7] we have, from the multiplier rule,

\[
\frac{d}{dt} \begin{pmatrix} f_{x_1'} x'_3 \\ x'_3 \\ x'_2 - x_1 \end{pmatrix} = \frac{d}{dt} F_{X'} = F_X = \begin{pmatrix} f_{x_1} x'_3 - x'_3 f_{x_2} \\ f_{x_2} \\ 0 \end{pmatrix}
\]

(2.3)

and

\[
F_{X',X'} = \begin{pmatrix} f_{x_1'x_1'} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.
\]

Hence \( \det F_{X',X'} = -f_{x_1'x_1'} \neq 0 \) so that, from [8], we have

**Theorem 2.1.** There exists a unique solution to (2.3) with the given boundary conditions.

It is instructive to reduce (2.3) to a second order problem in \( x_1(t) \). To do this we note we have

\[
x'_3(t) = x_1(t)
\]

(2.4)
from the third component of (2.3) and the third boundary condition for \( X(b) \), \((x'_2 - x_1)(b) = 0\). Hence

\[
x_2(t) = \int_a^t x_1(s) \, ds \quad \text{since} \quad x_2(a) = 0 \quad \text{and} \quad (2.5)
\]

\[
x'_3(t) = \int_a^t f x_2 \, ds + c \quad (2.6)
\]

from the second component of (2.3). But \( x'_3(b) = 0 \), by the second component condition of \( X(b) \), so that \( c = -\int_a^b f x_2 \, ds \) and hence

\[
x'_3 = -\int_t^b f x_2 \, ds. \quad (2.7)
\]

Finally,

\[
\frac{d}{dt}(f x') = f x_1 - x'_3 = f x_1 + \int_t^b f x_2 \, ds \quad (2.8)
\]

From the first component condition of (2.3). Thus, these results generalize the usual case when \( f \) is independent of \( x_2 = \hat{x} \). In addition, we have

**Theorem 2.2.** The solution to Problem (2.1) satisfies (2.8) and the boundary conditions in (2.1). If \( x_1(a) \) is not specified then \( f x'_1(a) = 0 \), while if \( x_1(b) \) is not specified then \( f x'_1(b) = 0 \).

We note that an equivalent expression for (2.8) is given in (3.1), below. The above are also the necessary conditions which allow us to construct solutions. In addition, we note that (2.8) is still a 2nd order differential equation, but it is not “ordinary” as it involves a (double) indefinite integral. In a simpler case it leads to a fourth order ODE as in Example 2.3, below.

As an exercise for the reader, we suggest obtaining the same results if we choose

\[
X(a) = \begin{pmatrix} x_a \\ 0 \\ * \\ 0 \end{pmatrix} \quad \text{and} \quad X(b) = \begin{pmatrix} x_b \\ * \\ 0 \end{pmatrix}
\]

We close this section by giving a simple example problem which leads to a fourth order ODE. Thus,

**Example 2.3.** Let \( f = \frac{1}{3}(x'^2 - \hat{x}^2) \), then \( f x'_1 = x'_1 \) and \( f x_2 = -x_2 \) so that \( x'_3 = \int_t^b x_2(s) \, ds \). The obvious calculations yield \( x''_1 = -\int_t^b x_2(s) \, ds \) and finally \( x^{(IV)}_1 - x_1 = 0 \). Thus, \( x = x_1 \) is in the span of \( S = \{e^{i t}, e^{-i t}, \sin t, \cos t\} \). The unique solution is found using the values \( x_1(a), x_1(b), \) and \( x''_1(b) = 0 \), since \( S \) is a linearly independent set of functions.

Finally, we note that the nonreduced \( f \) in (2.2) and (2.3) yields a more complicated second order differential equation than (2.8) which has a unique solution. However, Theorems 2.1 and 2.2 hold. In fact, (2.2) yields a (unique) numerical solution (see [5] or [6]) with an a priori error estimate of \( O(h^2) \). In the classical case where \( f \) is independent of \( x_2 \), the same situation holds for (2.8).
3. Variational Methods. The purpose of this section is to derive the basic result (2.8), once again. This time we will use an interesting new modification of more classical variational methods. We will see that what is required is new ideas involving \( \int_t^s \eta(s) \, ds \) where \( \eta(t) \) is the variation in \( x(t) \). Thus, we have

**Theorem 3.1.** If \( x(t) \) gives a minimum to Problem (1.1), then it satisfies

\[
\frac{d}{dt} f_x' = f_x + \int_b^t f_x'(s) \, ds ,
\]

(3.1)

and the boundary conditions of Theorem 2.2 hold.

*Proof.* Assume that \( x(t) \) yields a minimum for Problem (1.1) and \( \eta(t) \) is a piece-wise smooth function on \([a, b]\). Let \( x(t, \epsilon) = x(t) + \epsilon \eta(t) \) and note that

\[
I(\epsilon) = \int_a^b f(t, x(t, \epsilon), x'(t, \epsilon), \hat{x}(t, \epsilon)) \, dt
\]

is differential w.r.t. \( \epsilon \). Now \( I(0) \leq I(\epsilon) \) for \( \epsilon \) in a neighborhood of 0 implies

\[
0 = I'(0) = \int_a^b \left[ f_x \eta(t) + f_x' \eta'(t) + f_x' \int_a^t \eta(s) \, ds \right] dt
\]

(3.3)

\[
= \int_a^b \left[ f_x - \frac{d}{dt} f_x' + \int_t^b f_x'(s) \, ds \right] \eta(t) + f_x' \eta \bigg|_a^b .
\]

In the above, \( f_x, \) etc., is evaluated along \((t, x(t), x'(t), \hat{x}(t))\).

The above is achieved by noting that

\[
\frac{d}{dt} \left[ - \int_t^b f_x(s) \, ds \cdot \int_a^t \eta(s) \, ds \right] = f_x(t) \int_a^t \eta(s) \, ds - \eta(t) \int_t^b f_x(s) \, ds
\]

and then integrating both sides from \( t = a \) to \( t = b \).

Now if we assume that \( \eta(a) = \eta(b) = 0 \), then (3.1) follows from the Fundamental Lemma of the Calculus of Variations [5, p. 38]. The boundary conditions follow immediately by the usual arguments.

We note that Theorems 3.1 and 2.2 are the same. \( \square \)

**Example 3.2 (Revisited).** Result (3.1) becomes \( \frac{d}{dt} x' = -\int_t^b \hat{x} \, ds \) and hence \( x'' = \hat{x} \) or \( x^{iv} - x = 0 \), as before. The specific solution follows from the two given conditions and \( x''(b) = 0 \).

4. The Optimal Control Extension. The purpose of this section is to extend the previous theory to an optimal control setting. Thus, we join ideas in Section 2 with previous work by the author in [5] or [7] which showed that the classical optimal control problems with “solutions” by the Pontryagin Maximal Principle can be extended (along with additional constraints) and efficiently solved by methods as in Section 2.

Of note is the development of this section. We first show that our problem and solution include the classical Pontryagin conditions.
Our extension of the classical optimal control problem is

\[
\min \int_a^b f(t, x, u, \dot{x}) \, dt \\
\text{s.t. } x' = g(t, x, u, \dot{x}), \ x(a) = x_a, \ x(b) = x_b. \tag{4.1}
\]

We will get a more general Pontryagin Maximal Principle which includes the basic problem and results of the classical problem. That is, we will “solve” problem (4.1), but stop in the middle to obtain the classical result(s). In each case we assume \( f_{uu} \) and \( g_{uu} \) are linearly independent to impose nonsingularity in the sense of Bliss [1].

Thus, let \( x_1 = x, \ x'_2 = u, \ x'_3 \) be a multiplier (this is usually \( p \) of the classical theory). Also let \( x_4 = \dot{x} = \int_a^t x_1 \, ds \) and hence \( x'_4 = x_1, x_4(0) = 0 \). Finally, let \( x'_5 \) be the multiplier for the extension. Reformulating (4.1) we have

\[
\min \int_a^b F(t, X, X') \, dt, \ X(a) = \begin{pmatrix} x_a \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \ X(b) = \begin{pmatrix} x_b \\ \ast \\ \ast \\ \ast \\ \ast \end{pmatrix}, \tag{4.2}
\]

where \( F(t, X, X') = f(t, x_1, x'_2, x_4) + x'_3 [x'_1 - g(t, x_1, x'_2, x_4)] + x'_5 [x'_4 - x_1] \). Now

\[
\frac{d}{dt} \begin{pmatrix} x'_3 \\ x'_1 - g(t, x_1, x'_2, x_4) \\ x'_5 \\ x'_4 - x_1 \end{pmatrix} = d \frac{d}{dt} f_{x'} = F_x = \begin{pmatrix} f_{x_1} - x'_3 g_{x_1} \\ 0 \\ 0 \\ f_{x_4} - x'_5 g_{x_4} \end{pmatrix}. \tag{4.3}
\]

We make a brief detour to recall that the basic optimal control problem is as follows:

\[
\min \int_a^b f(t, x, u) \, dt \\
\text{s.t. } x' = g(t, x, u), \ x(a) = x_a, \ x(b) = x_b. \tag{4.4}
\]

Its “solution” is obtained as follows: let \( \mathcal{H}(t, x, u, p) = f(t, x, u) + p^T g(t, x, u) \), then

\[
x' = \mathcal{H}_p \tag{4.5a}
\]

\[
p' = -\mathcal{H}_x, \quad \text{and} \tag{4.5b}
\]

\[
\mathcal{H}_u = 0. \tag{4.5c}
\]

**Theorem 4.1.** Conditions (4.5) are included in (4.3) and (4.2). That is, if we ignore the \( x_1 \) and \( x_5 \) variables, we obtain (4.5).

**Proof.** Using the third component of (4.3) and the second component of \( X(b) \) we have

\[
x'_1 = g(t, x_1, x'_2, x'_4) \quad \text{which is (4.5a).} \tag{4.6a}
\]
Using the second component of (4.3) and the third component of $X(b)$ we have

$$f_{x_2} - x_2'g_{x_2} \equiv 0$$ which is (4.5c) with $-p = x_3'$.

(4.6b)

Finally, using the first component of (4.3) with $-p = x_3'$ we have

$$\frac{d}{dt}x_3' = f_{x_1} - x_3'g_{x_1}$$ which is (4.5b).

(4.6c)

Returning to (4.1), the more general problem, our first task is to check the non-singularity of (4.3). Thus,

$$f_{X \cdot X'} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & f_{x_1} - x_3'g_{x_1} - g_{x_2}' & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and, hence, $f_{X \cdot X'} = f_{x_2} - x_3'g_{x_2} = f_{uu} - x_3'g_{uu} \neq 0$. Thus,

**Theorem 4.2.** Equation (4.3) is a second order differential equation in five variables. This equation and the boundary conditions (4.2) give a unique solution to Problem 4.1. Thus, if (4.1) has a unique solution it is given by (4.3) and (4.2).

To complete the solution of Problem (4.1) we note that the fifth component of (4.3) and the fifth component of $X(b)$ implies

$$x_4' = x_1 \text{ and hence by the fourth component of } X(b),$$

$$x_4 = \int_a^t x_1(s) \, ds.$$  

(4.6d)

Finally, from the fourth component of (4.3) and the fourth component of $X(a)$, we have

$$x_5' = \int_a^t (f_{x_4} - x_3'g_{x_4})(s) \, ds + c \text{ or}$$

$$x_5' = \int_a^t (f_{x_4} - x_3'g_{x_4})(s) \, ds.$$  

(4.6e)

But, by the fourth component of $X(b)$ we have $x_5'(b) = 0$ and hence (also)

$$x_5'(t) = -\int_t^b (f_{x_4} - x_3'g_{x_4})(s) \, ds.$$  

5. Higher Derivatives. In this section we use our previous ideas to study

$$\min \int_a^b f(t, z(t), z'(t), z''(t)) \, dt$$

s.t. $f_{z''z''} > 0$, $z'(a) = A$, and $z'(b) = B$.  

and
We will see that the “s.t.” conditions are the natural conditions for this problem using our approach. This will explain and “clear up” the usual difficulties often encountered with these problems [2]. The extension of these ideas to the problem for \( f(t, x, x', \ldots, x^{(n)}) \) for \( n \geq 2 \) is immediate and left to the reader.

Our results in this section are motivated by the idea that Problem (1.1) contains three levels of differentiation but the variables and conditions must be chosen carefully. Motivated by Section 2 we set

\[ x_1(t) = z(t) \text{ and } x_2(t) = z'(t). \]  

Hence

\[ x_1'(t) = z'(t), \quad x_2'(t) = z''(t), \quad \text{and} \quad x_1' = x_2. \]  

Thus, we have

**Theorem 5.1.** If (5.1) has a unique solution then it satisfies

\[ \frac{d}{dt} f_{z''} = f_{z'} + \int_b^t f_z \, ds \]  

such that \( z'(a) = A \) and \( z'(b) = B \).

In addition, we note that

**Corollary 5.2.** Using (5.5) and (5.6), below, with (5.2), the algorithm in [5] or [6] leads to a numerical solution with an a priori, global error is \( O(h^2) \).

To justify these results and motivated by Section 2, we choose \( X = (x_1, x_2, x_3)^T \) and

\[ F(t, X, X') = f(t, x_1, x_2, x'_2) + x'_3(x'_1 - x_2) \]  

and

\[ X(a) = \begin{pmatrix} 0 \\ A \\ 0 \end{pmatrix}, \quad X(b) = \begin{pmatrix} * \\ B \\ * \end{pmatrix}. \]  

Now \( \frac{d}{dt} F_{X'} = F_X \) and hence

\[ \frac{d}{dt} \begin{pmatrix} x'_3 \\ f_{x'_2} \\ x'_1 - x_2 \end{pmatrix} = \begin{pmatrix} f_{x_1} \\ f_{x_2} - x'_3 \\ 0 \end{pmatrix}. \]  

As before, we note that

\[ F_{X'_X'} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & f_{x'_2 x'_2} & 0 \\ 1 & 0 & 0 \end{pmatrix} = -f_{x'_2 x'_2} \neq 0 \]  

so (5.4) has a unique solution.

Proceeding the (now) usual way we have

\[ x'_1 \equiv x_2 \text{ and thus } x_1(t) = \int_a^t x_2(s) \, ds, \]  

(5.8a)
Thus, (5.8c) is a second order ODE with double integral arguments. Differentiation yields a third order ODE with an integral argument. Under the best of conditions, where \( f \) is quadratic, it becomes a fourth order ODE which "agrees" with the classical results.

6. The \( g(t) \) Mean. The purpose of this section is to extend Problem (1.1) to the case where \( \dot{x} = g(t) \int_a^t x(t) \, dt \), and in particular, \( g(t) = \frac{1}{t-a} \), to obtain the “running mean” of \( x(t) \) as a new variable. In fact, we will derive the more general case. Hence, 

\[
\begin{align*}
\min \int_a^b f(t, x(t), x'(t), \overline{t}(t)) \, dt \\
\text{s.t. } f_{x'x'} > 0, \quad x(a) = x_a, \quad x(b) = x_b \\
\overline{t}(t) = g(t) \int_a^t x(t) \, dt.
\end{align*}
\]

We assume this problem has a unique solution while “smoothness” conditions will be obvious and assumed as needed. Following, from Section 2 we have

\[ x_2(t) = g(t) \int_a^t x_1 \, ds, \quad (6.2) \]

\[ x_2'(t) = g' \int_a^t x_1 \, ds + gx_1 = g \frac{x_2}{g} + gx_1, \quad (6.3) \]

\[ F = f(t, x_1, x_1', x_2) + x_3 \left[ x_2' - \frac{g'}{g} x_2 - gx_1 \right], \quad (6.4) \]

\[ \frac{d}{dt} \begin{pmatrix} f_{x_1'} \\ x_3' \\ x_2' - \frac{g'}{g} x_2 - gx_1 \end{pmatrix} = \begin{pmatrix} f_{x_1} - g x_3' \\ f_{x_2} - \frac{g'}{g} x_3' \\ 0 \end{pmatrix}, \quad (6.5) \]

\[ X(a) = \begin{pmatrix} x_a \\ 0 \\ 0 \end{pmatrix}, \quad X(b) = \begin{pmatrix} x_b \\ * \\ * \end{pmatrix}. \quad (6.6) \]

We note that we assume \( g(a) \) finite, so \( x_2(a) = 0 \). If \( g(t) = \frac{1}{t-a} \), then \( x_2(a) = 1 \), by l’Hospital’s Rule.

Hence \( x_2' = \frac{g'}{g} x_2 + gx_1 \) from the third component of (6.5) and \( X_3(b) \) in (6.6), and

\[ x_2(t) = g(t) \int_0^t x_1 \, ds \text{ from } X_2(a) \text{ in } (6.6). \quad (6.7a) \]
From \( x_3'' = f_{x_2} - \frac{g'}{g} x_3' \) we have \((g x_3')' = g x_3'' + g' x_3' = g (x_3'' + \frac{g'}{g} x_3') = g f_{x_2} \). Hence 
\[ g x_3' = \int_a^t g f_{x_2} ds + c_1 \]
or
\[ x_3' = \frac{1}{g} \left[ \int_a^t g f_{x_2} ds + c_1 \right] = -\frac{1}{g} \int_t^b g f_{x_2} ds \quad (6.7b) \]
since the second component of \( X(b) \) implies \( x_3'(b) = 0 \).

Finally, we have
\[ \frac{d}{dt} (f_{x_1'}) = f_{x_1} + \int_t^b g f_{x_2} ds. \quad (6.7c) \]

Note that all boundary conditions are used (accounted for). The third component of 
\( X(a) \) in (6.7) “normalize” \( x_3 \) as \( x_3' \) is unique, but \( x_3 \) is otherwise not. The values \( x_a \) and \( x_b \) along with the two starred conditions determine the solution for the second 
order equation (6.7).

Acknowledgement. The author wishes to acknowledge the thoroughness and 
patience of the referee for improving the quality of this paper.

REFERENCES

[1] G.A. Bliss, Lectures on the Calculus of Variations, Phoenix Science Series, University of 
equations, Utilitas Mathematica, 63 (2003), pp. 53–64.
[10] M. Kamien and N. Schwartz, Dynamic Optimization: The Calculus of Variations and Op-
p. 1814.