On the Foundations of Dynamic Programming

HOWARD OSBORN

Communicated by R. BELLMAN

1. Introduction. Let X_T be the class of all piecewise continuously differentiable functions x on $[0, T]$ such that $x(0) = 0$, let F be a function of two real variables satisfying conditions which will be specified in §2, and set

$$(1.1) \quad f(c, T) = \min_{\substack{x \in X_T \\ x(T) = c}} \int_0^T F(x(t), x'(t)) dt.$$

One of the basic problems of dynamic programming is to approximate the value $p(c, T)$ of the *optimal policy* or *slope function* at the point (c, T) in terms of F and the values $f(c_1, T - \epsilon)$, for some fixed $\epsilon > 0$ and a suitable range of values of c_1 . In this formulation p is the slope function of the field of extremals through $(0, 0)$; however, the considerations which follow can equally well be applied to other fields.

Heuristically, one may use BELLMAN'S principle of optimality as follows. For any $\epsilon > 0$,

$$(1.2) \quad \min_{\substack{x \in X_T \\ x(T) = c}} \int_0^T F(x(t), x'(t)) dt = \min_{\substack{x \in X_T \\ x(T) = c}} \left\{ f(x(T - \epsilon), T - \epsilon) + \int_{T-\epsilon}^T F(x(t), x'(t)) dt \right\},$$

and since $p(c, T) = x'(T)$ for the minimizing function x , one approximates the functions of X_T by setting $x(T - \epsilon) = c - \epsilon c' + o(\epsilon)$ and expands (1.2) in terms of ϵ to obtain

$$(1.3) \quad F(c, T) = \min_{c'} \{ f(c - \epsilon c', T - \epsilon) + \epsilon F(c, c') + o(\epsilon) \}.$$

Ignoring the $o(\epsilon)$, it appears plausible that if $p_*(c, T)$ is that value of c' which minimizes the expression

$$(1.4) \quad f(c - \epsilon c', T - \epsilon) + \epsilon F(c, c')$$

then

$$(1.5) \quad \lim_{\epsilon \rightarrow 0} p_*(c, T) = p(c, T).$$

This article was completed under the Research Grant NSF G5649.

Assuming that (1.5) is valid uniformly in c , one may approximate $f(c, T)$ to within $o(\epsilon)$ by the expression

$$(1.6) \quad f(c - \epsilon p_*(c, T), T - \epsilon) + \epsilon F(c, p_*(c, T))$$

for a suitable range of c 's to find $p(c, T + \epsilon)$ and $f(c, T + \epsilon)$ to within $o(1)$ and $o(\epsilon)$, respectively, for any given c . Similarly, given the approximation of $f(c, T + \epsilon)$ for a suitable range of c 's one may approximate $p(c, T + 2\epsilon)$ and $f(c, T + 2\epsilon)$, and so forth, thus obtaining the values of the slope function and the Hilbert integral at each point of the field domain. Many variations of this theme may be found in [1].

There are two objections to the preceding argument which cast doubt on (1.5). The first is that the class of all linear functions through (c, T) over the interval $[T - \epsilon, T]$ cannot replace the larger class of all piecewise continuously differentiable functions through (c, T) and over the same interval, and the second is that there is no apparent justification for removing $o(\epsilon)$ from the right-hand side of (1.3). Although the former objection can be met in part by another heuristic argument, which is presented in §4, there is no direct answer to the latter objection. It is therefore necessary to justify (1.5) by an entirely different argument.

In the present paper, (1.5) is proved by means of CARATHÉODORY'S approach to the classical calculus of variations, which is sketched in §2 for convenience. As a consequence of the proof itself, given in §3, it becomes clear that one can also approximate $p(c, T)$ by minimizing other expressions than (1.4). Two such expressions are obtained heuristically in §4 and §5, and the analogs of (1.5) are proved as in §3.

2. Carathéodory's development. We assume for convenience that F possesses continuous second derivatives and that it provides the well known sufficient conditions for a minimizing $x \in X_T$ for all values of c and c' occurring in the sequel. Specifically we assume that there exist unique functions p and q of c and t satisfying

$$(2.1) \quad dF_{c'}(c, p) = F_c(c, p) dt + q(dc - p dt)$$

such that the extremals x given by $dx(t)/dt = p(x(t), t)$ admit no envelopes in $[0, T]$ (the *strict Jacobi condition*), and we assume that

$$(2.2) \quad F(c, c') - F(c, p) \geq (c' - p)F_{c'}(c, p)$$

for any c and c' , where p is evaluated at (c, t) and where equality occurs only for $c' = p(c, t)$ (the *strict Weirstrass condition*).

From (2.1) it follows that $F(c, p)dt + F_{c'}(c, p)(dc - p dt)$ is an exact differential whose integral along the extremal from $(0, 0)$ to (c, T) is $f(c, T)$, hence that

$$(2.3) \quad f(c, T) = \int_{\Gamma} F(x, p(x, t)) dt + F_{c'}(x, p(x, t))(dx - p(x, t) dt)$$

along any path Γ joining $(0, 0)$ to (c, T) (the Hilbert *invariant integral*). Consequently f is continuously differentiable, and one obtains the *fundamental relation* of the calculus of variations

$$(2.4) \quad df(c, t) = F(c, p) dt + F_{c'}(c, p)(dc - p dt)$$

where p is evaluated at (c, t) , so that in particular

$$(2.5) \quad \begin{cases} F_{c'}(c, p) = f_c(c, T) \\ F(c, p) = f_T(c, T) + pf_c(c, T) \end{cases}$$

where p is evaluated at (c, T) . Substituting the latter relations into the Weierstrass condition (2.2) it follows that

$$(2.6) \quad F(c, c') - f_T(c, T) - c'f_c(c, T) \geq 0$$

for all c, c' , and T , with equality occurring only for $c' = p(c, T)$.

CARATHÉODORY regards (2.6) as a fundamental assumption in his development of the calculus of variations (see pp. 200–201 of [3]), and he uses it in place of the Weierstrass condition. We remark that by rewriting (1.1) as

$$(2.7) \quad \min_{x \in X_T} \left[\int_0^T F(x(t), x'(t)) dt - f(x(T), T) \right] = 0,$$

so that

$$(2.8) \quad \min_{x \in X_T} \left[\int_0^T F(x(t), x'(t)) - f_c(x(t), t)x'(t) - f_T(x(t), t) \right] dt = 0,$$

one can also obtain a heuristic proof of (2.6); it is difficult to proceed any further in this fashion, however.

3. The convergence proof. Suppose that Φ_* and Φ are continuously differentiable functions of c' such that

$$(3.1) \quad \lim_{\epsilon \rightarrow 0} \frac{\partial}{\partial c'} \Phi_*(c') = \frac{\partial}{\partial c'} \Phi(c')$$

uniformly in some closed neighborhood N which contains a unique minimum p of Φ . Then clearly Φ_* attains a minimum value for some unique p_* , provided that ϵ is sufficiently small, and

$$(3.2) \quad \lim_{\epsilon \rightarrow 0} p_* = p.$$

Furthermore, if (3.1) is uniform in some parameters c and T it is clear that the same is true of (3.2).

To prove (1.5) we first note that, for any $\epsilon > 0$, (1.4) and the expression

$$(3.3) \quad F(c, c') = \frac{f(c, T) - f(c - \epsilon c', T - \epsilon)}{\epsilon}$$

are minimized by the same c' , which was denoted by $p_*(c, T)$ in §1. If $\Phi_*(c')$ denotes the expression in (3.3) and

$$(3.4) \quad \Phi(c') = F(c, c') - f_c(c, T)c' - f_T(c, T),$$

then (3.1) is satisfied by the continuity of f_c . But $\Phi(c')$ is uniquely minimized by the value $p(c, T)$ of the slope function at (c, T) , by (2.6), so that we obtain

Theorem. *Let F satisfy the strict Jacobi and Weierstrass conditions for the field of extremals through $(0, 0)$ with slope function p , and let $f(c, T)$ denote the value*

$$\min_{\substack{x \in X_T \\ x(T) = c}} \int_0^T F(x(t), x'(t)) dt$$

of the Hilbert integral along paths joining $(0, 0)$ to (c, T) . Then for sufficiently small $\epsilon > 0$ there exists a unique value $p_(c, T)$ of c' which minimizes $f(c - \epsilon c', T - \epsilon) + \epsilon F(c, c')$, and*

$$\lim_{\epsilon \rightarrow 0} p_*(c, T) = p(c, T).$$

It is worth noting that $p(c, T)$ need not be an interior minimum of $\Phi(c')$.

4. The ϵ^2 terms. If $x(T - \epsilon)$ is replaced by $c - \epsilon c' + \frac{1}{2}\epsilon^2 c''$ in the heuristic derivation of (1.4), then (1.4) is replaced by the expression

$$(4.1) \quad f(c - \epsilon c' + \frac{1}{2}\epsilon^2 c'') + \epsilon F(c, c') - \frac{1}{2}\epsilon^2 [F_c(c, c')c' + F_{c'}(c, c')c''],$$

whose minimum with respect to c' and c'' leads to an approximation of $p(c, T)$ to within $o(\epsilon^2)$, as one easily shows by the method of §3.

It is interesting to note that if one formally expands (4.1) in powers of ϵ and differentiates the coefficients of ϵ and ϵ^2 with respect to c' , one obtains the conditions

$$(4.2) \quad f_c = F_{c'},$$

and

$$(4.3) \quad 2(f_{cc'}c' + f_{cT}) = F_{cc'}c' + F_{c'c'}c'' + F_c,$$

respectively, for the minimizing c' . But (4.2) is just the fundamental relation (2.4), whose material derivative is

$$(4.4) \quad f_{cc'}c' + f_{cT} = F_{c'c'}c' + F_{c'c'}c'',$$

and, by means of (4.4), one observes that (4.3) is just the Euler variational equation.

The coefficient of ϵ in (4.1) does not depend on c'' , and if one differentiates the coefficient of ϵ^2 with respect to c'' , the minimizing condition is just (4.2) again. Thus one does not obtain any new identities by introducing c'' . Similarly,

by introducing additional new terms in the Taylor series of $x(T - \epsilon)$, one never obtains identities which cannot also be obtained by considering the approximation $c - \epsilon c'$. This is the heuristic justification for replacing the class of piecewise continuously differentiable functions over $[T - \epsilon, T]$ in §1 by the class of linear functions through (c, T) .

5. A new algorithm. Suppose that a field of extremals through (c, T) has been constructed over the interval $[T - \epsilon, T]$. Then if $x_{c'}$ denotes that extremal which passes through (c, T) with slope c' ,

$$(5.1) \quad \int_{T-\epsilon}^T F(x_{c'}(t), x'_{c'}(t)) dt = \int_{\Gamma} [F(x, p) - F_{c'}(x, p)p] dt + F_{c'}(x, p) dx$$

where the slope-function p is evaluated at (x, t) and where Γ is any path in the field domain which joins $(T - \epsilon, x_{c'}(T - \epsilon))$ with (T, c) . In particular, we may suppose that Γ consists of two line-segments Γ_x and Γ_t which join $(T - \epsilon, x_{c'}(T - \epsilon))$ with $(T - \epsilon, c)$ and $(T - \epsilon, c)$ with (T, c) , respectively. Then the principle of optimality implies that

$$(5.2) \quad \begin{aligned} f(c, T) &= \min_{c'} \left\{ f(x_{c'}(T - \epsilon), T - \epsilon) + \int_{T-\epsilon}^T F(x_{c'}(t), x'_{c'}(t)) dt \right\} \\ &= \min_{c'} \left\{ f(x_{c'}(T - \epsilon), T - \epsilon) + \int_{\Gamma_t} (F - F_{c'} p) dt + \int_{\Gamma_x} F_{c'} dx \right\}, \end{aligned}$$

and since the first integral on the right-hand side of (5.2) is independent of c' it follows that $p(c, T)$ is that value of c' which minimizes

$$(5.3) \quad f(x_{c'}(T - \epsilon), T - \epsilon) + \int_{x=c}^{x_{c'}(T-\epsilon)} F_{c'}(x, p(x, T - \epsilon)) dx,$$

for any ϵ . Proceeding heuristically as in §1, we note that we may write $x_v(T - \epsilon) = c - \epsilon y + o(\epsilon)$, and, in particular, $x_0(T - \epsilon) = c + o(\epsilon)$, so that the integral in (5.3) may be written in terms of y as $\epsilon \int_{y=0}^{c'} F_{c'}(c - \epsilon y, y) dy + o(\epsilon)$. Hence one expects that that value $p^*_\epsilon(c, T)$ of c' which minimizes

$$(5.4) \quad f(c - c'\epsilon, T - \epsilon) + \epsilon \int_{y=0}^{c'} F_{c'}(c - \epsilon y, y) dy$$

satisfies

$$(5.5) \quad \lim_{\epsilon \rightarrow 0} p^*_\epsilon(c, T) = p(c, T).$$

In order to prove (5.5), we note that, for any $\epsilon > 0$, $p^*_\epsilon(c, T)$ also minimizes the expression

$$(5.6) \quad \int_{y=0}^{c'} F_{c'}(c - \epsilon y, y) dy - \frac{f(c, T) - f(c - \epsilon c', T - \epsilon)}{\epsilon},$$

which we again denote by $\Phi_*(c')$. Then if we set

$$(5.7) \quad \Phi(c') = F(c, c') - f_c(c, T)c' - f_T(c, T)$$

as in (3.2), the conditions stated at the beginning of §3 are again satisfied due to the continuity of $F_{c'}$ and f_c , and the proof of (5.5) follows as in §3.

BIBLIOGRAPHY

- [1] BELLMAN, R., *Dynamic Programming*, Princeton University Press, Princeton, New Jersey, 1957.
- [2] BLISS, G. A., *Calculus of Variations*, Open Court, Chicago, 1925.
- [3] CARATHÉODORY, C., *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*, B. G. Teubner, Leipzig, 1935.

The University of Illinois
Urbana, Illinois
and
The RAND Corporation
Santa Monica, California