A NEW DEFINITION OF THE INTEGRAL FOR NON-PARAMETRIC PROBLEMS IN THE CALCULUS OF VARIATIONS

BY

JAMES SERRIN

University of Minnesota

In this paper we shall be concerned with certain multiple integrals which arise in the calculus of variations, namely those of the form

$$I[u] = \int_R F(x,y,u,p,q)\,dA.$$

Here R denotes an open bounded region of the plane, the integrand F is subject to certain conditions which will be stated later, and $dA = dx\,dy$.

The calculus of variations is concerned with minimizing the integral $I[u]$ in some class of functions for which $I[u]$ is defined. It is natural to want this class to be as large as possible, since the existence of a minimizing function is usually more easily established in an extended class; there are also aesthetic reasons for considering a broad class of admissible functions, as may be seen from the theory of surface area. For reasons such as these, we have introduced in the preceding paper two rather extensive classes of admissible functions, the members of which in one instance need not even be everywhere continuous. Now the introduction of these classes entailed in both cases a *definition* of the integral $I[u]$: the purpose of this paper is to show that these definitions are logically consistent, that is, supply the correct (natural) value whenever u is continuously differentiable.

From another point of view, the paper may be considered as simply proposing a new definition of the integral $I[u]$, somewhat analogous to the generalized area in the theory of surface area. In order to make the work more accessible to the general reader, we have included in section 1 a complete statement of results, so that the paper is entirely self-contained. Section 2 is devoted to a generalization of a lemma of Tonelli, which will be useful in the sequel, while the main proofs are carried out in sections 3 to 5. It is of related interest to know whether the class of functions considered here includes the usual functions of importance in the calculus of variations (i.e. Lipschitzian functions, absolutely continuous functions, etc.). We shall address some remarks to this subject at the close of the paper.

The integral $I_c[u]$ which we define below may be considered in an alternative way if one so wishes, namely, it can be thought of as arising from the well-known Frechet extension of a lower semi-continuous operator. Looked at this way, Theorem 2 then becomes the necessary verification of the lower semi-continuity property of $I[u]$. Although this point of view does not seem to be as fundamental as the one actually adopted, it does serve to connect the present paper with other investigations in the calculus of variations. In particular, our Theorem 2 is closely related with certain work of Tonelli [8] and Cinquini [3] on the lower semi-continuity of integrals, though it appears that the proof methods which we use are somewhat simpler than in those papers.

1. Hypotheses and definitions

We assume that the integrand $F(x,y,u,p,q)$ is defined and continuous for all $(x,y) \in R$ and all values of (u,p,q). Furthermore, we suppose that the partial derivatives F_p and F_q exist and are continuous, and that

$$F(x,y,u,p,q) \geqslant 0 \tag{1}$$

and
$$E(x,y,u,p,q;P,Q) \geqslant 0, \quad (P,Q) \neq (p,q). \tag{2}$$

The main result of the paper (Theorem 2) requires a further condition on the integrand. We consider two possibilities:

§ 1. *The function F has continuous partial derivatives F_{pp}, F_{pq}, F_{qq}, F_{px},* and F_{qy}.

§ 2. *Condition* (2) *holds with the equality sign excluded.*

Though these hypotheses can be weakened somewhat (see section 4), it is still not known whether the conclusion of Theorem 2 holds merely under the assumptions (1) and (2).

We turn now to the definition of $I[u]$, which forms the heart of the paper. Actually, we shall give several definitions, depending on the nature of the function $u = u(x,y)$.

1. *u is continuously differentiable in R.* The integral $I[u]$ is defined in the obvious way, as the improper Riemann integral obtained by exhaustion of the region R; that is,

$$I[u] = \lim_{S \to R} I[u, S],$$

S here denoting a proper subregion of R.(¹)

(¹) It would also be possible to define $I[u]$ as the Lebesgue integral of $F(x, y, u, u_x, u_y)$ over R, but we prefer the present method on account of its greater simplicity.

Of course, if u has bounded derivatives in the whole of R it is obviously not necessary to use even the concept of improper integral.

2. *u is continuous in R.* Consider a sequence $\{u_n\}$ of continuously differentiable functions, each function u_n being defined in a closed subregion R_n of R. Suppose that

$$R_n \to R, \qquad \mathrm{Max}\,|u - u_n| \to 0$$

as n tends to infinity. Then we set

$$I_C[u] = \text{g.l.b.}\ \liminf I[u_n, R_n],$$

where the g.l.b. is taken over all sequences of the type described above. $I_C[u]$ is clearly well-defined, since the class of sequences in question is non-empty.

3. *u is measurable in R.* The only change from the preceding definition is that uniform convergence is replaced by almost everywhere convergence. The resulting integral will be denoted by $I_M[u]$.

The second and third definitions are generalizations of well-known definitions of surface area in the non-parametric case (cf. references [1, 2, 4, 5]). This fact suggests that a theory of the integrals I_C and I_M be developed comparable to the theory of non-parametric surface area. The consistency theorem mentioned in the introduction is one step in such a program. Before stating this result, let us note the very agreeable semi-continuity properties which I_C and I_M possess as immediate consequences of their definition.

THEOREM 1. *The integral $I_C[u]$ is lower semi-continuous in the class of continuous functions with the uniform open topology. More precisely, if u_n, $n = 1, 2, \ldots$, is continuous in an open region R_n and*

$$R_n \to R, \qquad \mathrm{Max}\,|u - u_n| \to 0$$

as n tends to infinity, then

$$I_C[u] \leqslant \liminf I_C[u_n, R_n].$$

A similar result holds for the integral $I_M[u]$; thus, if u_n is measurable in an open region R_n and $R_n \to R$, $u_n \to u$ almost everywhere, then $I_M[u] \leqslant \liminf I_M[u_n, R_n]$.

The main result of the paper is that the integrals I_C and I_M represent successive extensions of $I[u]$. This may be stated in the form of two theorems:

THEOREM 2. *If either hypothesis $\mathfrak{H}$ 1 or $\mathfrak{H}$ 2 holds, then $I_C[u] = I[u]$ for each continuously differentiable function u in R.*

THEOREM 3. *If u is continuous in R, then $I_M[u] = I_C[u]$.*

The proof of Theorem 2 will be given in sections 3 and 4 of the paper, while the proof of Theorem 3 will be given in section 5. We note that throughout the paper the restriction to two independent variables is purely notational.

2. A lemma of Tonelli

A one-dimensional form of the following lemma appears in Tonelli's *Fondamenti di calcolo delle variazioni*, the two-dimensional form in his *Acta Mathematica* paper (but with a less explicit estimate). The interesting thing about the lemma, and the fact which makes it useful in the present application, is that the estimate (3) is obtained without using any differentiability properties of Q with respect to u.

LEMMA 1 (Tonelli). *Let $u = u(x,y)$ be a continuously differentiable function defined in a closed region S with smooth boundary Σ. We define*

$$J[u] = \int_S Q(x,y,u)\, p\, dA,$$

where Q and its partial derivative $\partial Q/\partial x$ are continuous functions. Then

$$|J[u] - J[v]| \leqslant M \oint_\Sigma |u-v|\, ds + M' \int_S |u-v|\, dA, \tag{3}$$

where $M = \operatorname{Max}|Q|$, $M' = \operatorname{Max}|\partial Q/\partial x|$.

Proof. Suppose first that $u \neq v$ throughout S, say $v > u$. Let S' denote the region in (x,y,z) space defined by the inequality

$$u(x,y) \leqslant z \leqslant v(x,y), \quad (x,y) \in S.$$

Applying the divergence theorem to this region we find

$$\iiint_{S'} \frac{\partial Q}{\partial x}\, dx\, dy\, dz = J[v] - J[u] + \iint_{\Sigma'} Q\, n_x\, da,$$

where Σ' denotes the lateral boundary of S', and n_x is the x-component of the unit outer normal vector to Σ'. The required estimate follows at once.

To prove (3) in the general case we simply approximate u and v by polynomials u_n and v_n. The preceding proof applies in each subregion S_i of S in which $u_n \neq v_n$; by adding over these regions we get

$$\begin{aligned} |J[u_n] - J[v_n]| &\leqslant \sum_i |J[u_n, S_i] - J[v_n, S_i]| \\ &\leqslant M \oint_\Sigma |u_n - v_n|\, ds + M' \int_S |u_n - v_n|\, dA. \end{aligned}$$

Letting n tend to infinity yields the inequality (3).

It should be observed that the bounds M and M' appearing in (3) can be given the more explicit form

$$M = \operatorname*{Max}_{\Sigma^*} |Q(x,y,z)|, \qquad M' = \operatorname*{Max}_{S^*} \left| \frac{\partial Q}{\partial x}(x,y,z) \right|,$$

where S^* denotes the set in (x,y,z)-space defined by the inequality

$$\mathrm{Min}[u(x,y), v(x,y)] \leqslant z \leqslant \mathrm{Max}[u(x,y), v(x,y)], \quad (x,y) \in S,$$

and Σ^* denotes the lateral boundary of S^*. This remark will be of importance in the sequel.

3. Proof of Theorem 2

Consider a function $u = u(x,y)$ which is continuously differentiable in R. Since u can be considered as the uniform limit of a sequence each of whose members is u itself, we obviously have $I_C[u] \leqslant I[u]$. We proceed now to prove the oppsite inequality. This proof will be carried through first under hypothesis $\mathfrak{H}$ 1, namely that the derivatives F_{pp}, F_{pq}, F_{qq}, F_{px}, and F_{qy}, exist and are continuous (the method is due to Tonelli).

Let $\varepsilon > 0$ be arbitrary, and let S denote a fixed closed subregion of R. We approximate u in S by a polynomial w such that

$$|u - w|, \; |\nabla u - \nabla w| < \varepsilon.$$

For the remainder of the proof we set $p = u_x$, $q = u_y$, $p' = w_x$, $q' = w_y$. Evidently p' and q' are bounded in S independent of ε, say

$$|p'|, \; |q'| < K \quad \text{in } S. \tag{4}$$

Now let U be any continuously differentiable function in S. Using the definition of E, we have the identity

$$F(U,P,Q) = F(U,p',q') + (P-p)F_p(U,p',q') + \\ + (Q-q)F_q(U,p',q') + E(U,p',q';P,Q), \tag{5}$$

in which the arguments x,y have been uniformly suppressed. From this equality follows

$$E(u,p',q';p,q) \leqslant 4\,\varepsilon M, \quad (x,y) \varepsilon S, \tag{6}$$

where M is an upper bound for $|F_p|$ and $|F_q|$ for the arguments (x,y,u,p',q'). In virtue of (4) it is clear that M can be chosen to depend only on S (and of course on the given function u). Also from (5) and (6) follows

$$
\begin{aligned}
&F(U,P,Q) - F(u,p,q) \\
&\geqslant F(U,p',q') + (P-p')F_p(U,p',q') + (Q-q')F_q(U,p',q') - \\
&- F(u,p',q') - (p-p')F_p(u,p',q') - (q-q')F_q(u,p',q') - 4\varepsilon M. \qquad (7)
\end{aligned}
$$

Now because of the assumption made concerning the derivatives of F, it follows that the functions

$$
\begin{aligned}
\mathcal{Q}(x,y,z) &= F_p(x,y,z,p'(x,y),q'(x,y)) \\
\mathcal{R}(x,y,z) &= F_q(x,y,z,p'(x,y),q'(x,y))
\end{aligned}
$$

have continuous derivatives $\partial \mathcal{Q}/\partial x$ and $\partial \mathcal{R}/\partial y$. Hence if we integrate (7) over S, and use Tonelli's lemma (section 2) and the continuity of F, F_p, and F_q, the termwise differences in the second and third lines of (7) each contribute a term which tends to zero with Max $|U-u|$. Thus we have

$$I[U,S] - I[u,S] \geqslant -\varepsilon_1 - 4\varepsilon M \text{ Area } S,$$

where $\varepsilon_1 \to 0$ as $\operatorname{Max}|U-u| \to 0$.

The proof can now be completed easily, subject of course to the differentiability conditions on F. Let $\{u_n\}$ be a sequence of continuously differentiable functions tending to u, as in the definition of $I_C[u]$. We may assume by the diagonal process that $I[u_n, R_n] \to I_C[u]$. This being the case we can choose n so that, simultaneously,

$$I[u_n,S] - I[u,S] \geqslant -\varepsilon - 4\varepsilon M \text{ Area } S$$

and

$$I[u_n,S] \leqslant I[u_n,R_n] \leqslant I_C[u] + \varepsilon.$$

Since ε is arbitrary, these two inequalities together imply

$$I[u,S] \leqslant I_C[u],$$

and the required inequality follows at once.

4. Proof of Theorem 2 (concluded)

In this section we shall prove the inequality $I[u] \leqslant I_C[u]$ under the alternate hypothesis $\mathfrak{H}$ 2. A simple lemma is needed first.

LEMMA 2. *Suppose that $E > 0$ for all $(P,Q) \neq (p,q)$, and let $\mathcal{D}$ denote a bounded set of values (x,y,u,p,q). Then there exists a constant $\varkappa > 0$ such that*

$$E \geqslant \varkappa X, \qquad X = \sqrt{(P-p)^2 + (Q-q)^2}, \qquad (8)$$

for all $(x,y,u,p,q) \in \mathcal{D}$ and all $X \geqslant 1$.

Proof. Since the figurative surface is convex (see, for example the lemma in section 9 of the preceding paper), it is clear that (8) will hold with

$$\varkappa = \underset{\mathcal{D}}{\text{g.l.b.}} \underset{X=1}{\text{Min}}\, E(x,y,u,p,q;\, P,Q).$$

The lemma being proved, now let S be a fixed closed subregion of R with smooth boundary, and let ε be an arbitrary positive number. We observe that

$$F(U,P,Q) - F(U,p,q) = (P-p)\,F_p(U,p,q) + (Q-q)\,F_q(U,p,q) + E(U,p,q;P,Q),$$

whence for $(x,y) \in S$,

$$\begin{aligned} &F(U,P,Q) - F(u,p,q) \\ &\qquad \geqslant (P-p)\,F_p(u,p,q) + (Q-q)\,F_q(u,p,q) + E(U,p,q;P,Q) - \varepsilon_2 X - \varepsilon_3, \end{aligned}$$

where ε_2 and ε_3 are positive numbers which tend to zero with $\text{Max}\,|U-u|$. Now the arguments (x,y,U,p,q) of E belong to some bounded set $\mathcal{D}$, so that by Lemma 2 there exists a constant $\varkappa > 0$ such that

$$E \geqslant \varkappa X \quad \text{if} \quad X \geqslant 1, \quad (x,y) \in S.$$

Moreover, the terms $F_p(u,p,q)$ and $F_q(u,p,q)$ can be approximated by polynomials A and B such that

$$|F_p - A|,\ |F_q - B| < \varepsilon, \quad (x,y) \in S.$$

It follows that, for $(x,y) \in S$,

$$F(U,P,Q) - F(u,p,q) \geqslant G - (P-p)A - (Q-q)B - \varepsilon_3,$$

where

$$G = \begin{cases} -2\varepsilon - \varepsilon_2 & \text{if } X < 1, \\ (\varkappa - 2\varepsilon - \varepsilon_2)\,X & \text{if } X \geqslant 1. \end{cases}$$

Let us assume from here on that $\varepsilon < \varkappa/4$ and that $\text{Max}\,|U-u|$ is so small that $\varepsilon_2 < \varkappa/2$. Then $G > 0$ for $X \geqslant 1$, and we find after an integration by parts that

$$I[U,S] - I[u,S] \geqslant -(2\varepsilon + \varepsilon_2)\,\text{Area}\, S - \varepsilon_4,$$

where ε_4 tends to zero with $\text{Max}\,|U-u|$. The remainder of the argument is exactly as in section 3, and Theorem 2 is completely proved.

Remark. It will be seen that hypothesis $\mathfrak{H}$ 2 was used simply to establish Lemma 2. Thus Theorem 2 remains true if instead of $\mathfrak{H}$ 2 we require only the following condition.

$\mathfrak{H}$ 2'. *For each bounded set $\mathcal{D}$ in the five-dimensional (x,y,u,p,q) space there exists a positive number X_0 such that $E > 0$ for all $(x,y,u,p,q) \in \mathcal{D}$ and $\sqrt{(P-p)^2 + (Q-q)^2} > X_0$.*

The differentiability condition $\mathfrak{H}$ 1 may also be lightened in certain cases, specifically whenever the integrand can be approximated by a twice differentiable function which satisfies the convexity condition (2). At first glance this may appear possible in all cases, but upon considerable reflection I have come to the conclusion that only certain integrands are amenable to this treatment, and that a general approximation procedure is not a simple matter after all.

5. Proof of Theorem 3

Let $u = u(x,y)$ be continuous in R. Since any sequence $\{u_n\}$ allowable in the definition of $I_C[u]$ is also allowable in the definition of $I_M[u]$, we obviously have $I_M[u] \leqslant I_C[u]$. We proceed now to prove the opposite inequality

$$I_C[u] \leqslant I_M[u]. \tag{9}$$

Let $\{u_n\}$ be a sequence of continuously differentiable functions (each function u_n being defined in a closed subregion R_n of R) with the properties

$$R_n \to R, \quad u_n \to u \text{ almost everywhere in } R,$$

and

$$I[u_n, R_n] \to I_M[u]. \tag{10}$$

The existence of such a sequence is guaranteed by the diagonal process and the definition of $I_M[u]$. Our method of proving (9) is to construct from the sequence $\{u_n\}$ another sequence $\{v_n\}$ of piecewise continuously differentiable functions such that

$$\operatorname{Max} |u - v_n| \to 0 \tag{11}$$

and

$$I[v_n, R_n] \leqslant I_M[u] + \varepsilon, \tag{12}$$

where ε is an arbitrary positive number. Once this is done, we can conclude that

$$I_C[u] \leqslant \liminf I[v_n, R_n] \leqslant I_M[u] + \varepsilon,$$

and (9) follows at once.

In order to construct the individual function v_n we first choose in R_n two polynomials w' and w'' satisfying

$$\left| w' - u - \frac{2}{n} \right| < \frac{1}{n}, \qquad \left| w'' - u + \frac{2}{n} \right| < \frac{1}{n}.$$

Let M be an upper bound for the functions $F(x,y,w',w'_x,w'_y)$ and $F(x,y,w'',w''_x,w''_y)$, for $(x,y) \in R_n$. Now according to a theorem of Egoroff $u_n \to u$ almost uniformly,[1] hence there exists a set C in R_n with the properties

$$\text{Measure } C \leqslant \varepsilon/2M \tag{13}$$

and

$$\lim_{m\to\infty} (\max_{R'_n} |u_m - u|) = 0, \quad R'_n \equiv R_n - C. \tag{14}$$

From (1), (10), and (14) follows the existence of an integer k such that

$$I[u_k, R_n] \leqslant I_M[u] + \tfrac{1}{2}\varepsilon, \qquad \max_{R'_n} |u_k - u| < 1/n. \tag{15}$$

We may assume without loss of generality that u_k is a polynomial. The function v_n is now defined by

$$v_n = \begin{cases} w' & \text{if } u_k > w', \\ w'' & \text{if } u_k < w'', \\ u_k & \text{otherwise.} \end{cases}$$

Clearly v_n is piecewise continuously differentiable, and can be substituted into the integral I.

Let C' and C'' denote, respectively, the point sets where $u_k > w'$ and $u_k < w''$. Then since $w'' < u - 1/n$ and $u + 1/n < w'$ one sees from $(15)_2$ that

$$C' + C'' \subset C. \tag{16}$$

Moreover, since $|w' - u|$ and $|w'' - u|$ are each less than $3/n$, we have $|v_n - u| < 3/n$, so that (11) is satisfied. Finally

$$\begin{aligned} I[v_n, R_n] &= I[w', C''] + I[w'', C''] + I[u_k, R_n - C' - C''] \\ &\leqslant M \cdot \text{Area}\,(C' + C'') + I[u_k, R_n] \\ &\leqslant I_M[u] + \varepsilon, \end{aligned}$$

using (16), (13), and $(15)_1$. This completes the proof.

6. Final remarks

Let $u = u(x,y)$ be a Lipschitzian function in R, that is, suppose that for each proper subregion B of R there exists a constant $M = M(B)$ such that the inequality

$$|u(x,y) - u(x_0,y_0)| < M(|x - x_0| + |y - y_0|)$$

---

[2] A sequence of functions ψ_n is said to converge almost uniformly to a function ψ in case for every positive δ there exists a set C such that Measure $C \leqslant \delta$ and $\lim \psi_n = \psi$ uniformly in the complement of C.

holds for (x,y), $(x_0,y_0) \in B$. Then the derivatives u_x and u_y exist almost everywhere and one may define the (Lebesgue) integral

$$I_L[u] = \int_R F(x,y,u,u_x,u_y)\,dA.$$

The question now arises, is $I_C[u] = I_L[u]$?

It is known that u can be approximated in any closed subregion R_n by continuously differentiable functions w such that

$$|\nabla w| < M(R_n) + 1, \qquad |u - w| \to 0,$$

and $\nabla w \to \nabla u$ almost everywhere. Consequently there exists a function w_n such that

$$|I[w_n, R_n] - I_L[u, R_n]| < 1/n,$$

and therefore

$$I_C[u] \leqslant \liminf I[w_n, R_n] = I_L[u].$$

The opposite inequality can be proved by the methods of sections 3 and 4. Thus it is proved that $I_C[u] = I_L[u]$; in other words, $I_C[u]$ supplies the "natural" value of the integral even in the class of Lipschitzian functions.

If u is absolutely continuous in the sense of Tonelli (cf. [8]), then $I_L[u]$ is still well-defined, and again it can be shown that $I_C[u] = I_L[u]$. The argument is similar to that above, though we shall omit the details (cf. in particular, [8], §§ 12, 13). In summary, *any class of admissible functions in the calculus of variations based on the integral $I_C[u]$ constitutes an extension of the corresponding class based on Lipschitzian functions or even absolutely continuous functions.*

References

[1]. L. Cesari, *Surface Area*. Ann. of Math. Studies No. 35, Princeton University Press, 1955. Especially pp. 21–24 and Appendix B.

[2]. ——, Sulle funzioni a variazione limitata. *Ann. Scuola Norm. Sup. Pisa* (2), 5 (1936), 299–313.

[3]. S. Cinquini, Condizioni sufficienti per la semicontinuitá nel calcolo delle variazioni. *Ann. Scuola Norm. Sup. Pisa* (2), 2 (1933), 41–58.

[4]. C. Goffman, Lower semi-continuity and area functionals. *Rend. Circ. Mat. Palermo* (2), 2 (1953), 203–235.

[5]. T. Radó, *Length and Area*. Amer. Math. Soc. Coll. Publ., 30, 1946. Especially chapter V.2.

[6]. J. Serrin, On a fundamental theorem of the calculus of variations. Preceding in *this Journal*, 1–22.

[7]. L. Tonelli, *Fondamenti di calcolo delle variazione*. Bologna, 1922. Especially vol. I, 385–392.

[8]. —— Sur la semi-continuité des intégrales doubles du calcul des variations. *Acta Math.*, 53 (1929), 325–346.