A MULTIPLIER RULE IN ABSTRACT SPACES*

HERMAN H. GOLDSTINE

The Lagrange multiplier rule has been generalized to certain non-calculus of variations problems by Graves,† Hahn,‡ and the author.§ Moreover a very general problem was formulated by L. A. Luster-nik.‖ However his work seems to rest upon a theorem which is stated without proof and which the author is unable to verify. There are also certain other difficulties with his proof. The problem herein presented is so formulated that the problem of Bolza in the calculus of variations, the problems treated by Graves, Hahn, and the author, and the problem of minimizing a functional defined upon an arbitrary Banach space subject to very general numerically-valued side conditions, and numerous other examples are included as special cases.

The proof proceeds along lines which are essentially generalizations of the methods of the calculus of variations. This demonstration is made possible by the very powerful implicit function theorems of Hildebrandt and Graves¶ and yields analogs of the transversality condition and of the Euler-Lagrange equations. Some instances which explain the number of linear spaces involved are given in the concluding section.

1. **Definitions and assumptions.** In order to obtain our statement and proof of the multiplier rule it is convenient to refer to five normed linear spaces $\mathfrak{M}, \mathfrak{N}, \mathfrak{X}, \mathfrak{U}, \mathfrak{V}$, of which $\mathfrak{M}, \mathfrak{U}$, and $\mathfrak{V}$ are assumed to be complete. It is further supposed that $\mathfrak{M}_0, \mathfrak{N}_0$, and $\mathfrak{X}_0$ are regions of their respective spaces. We shall then be concerned with the following basis:

A. f and ψ_k, $(k = 1, \cdots, q)$, where q is an arbitrary integer, are real-valued functions defined on the composite* $(\mathfrak{X}_0, \mathfrak{M}_0)$ of $\mathfrak{X}_0$ and $\mathfrak{M}_0$, each of which has a first differential at some fixed point $(\bar{x}_0, \bar{\mu}_0)$. Moreover the functionals ψ_k, $(k = 1, \cdots, q)$, vanish at this point $(\bar{x}_0, \bar{\mu}_0)$.

B. $u = \phi_1(x_0, \mu_0, \nu_0)$ is an operation of class C' defined on $(\mathfrak{X}_0, \mathfrak{M}_0, \mathfrak{N}_0)$, and having its functional values in $\mathfrak{U}$; and $\nu_0 = n(x_0, \mu_0)$ is defined over $(\mathfrak{X}_0, \mathfrak{M}_0)$, has its values in $\mathfrak{N}_0$, is of class C', and is such that $\phi_1[\bar{x}_0, \bar{\mu}_0, n(\bar{x}_0, \bar{\mu}_0)] = 0_u$, the zero element in the space $\mathfrak{U}$. (Analogous notations will be used for the zero points in the various spaces treated.)

C. There is an operation $v = \phi_2(x_0, \mu_0, \nu_0)$ of class C', which is defined on a neighborhood of the point $(\bar{x}_0, \bar{\mu}_0, \bar{\nu}_0)$, where $\bar{\nu}_0 = n(\bar{x}_0, \bar{\mu}_0)$, and has its functional values in $\mathfrak{V}$, such that the linear and continuous function of μ, $(d_\mu\phi_1(x_0, \mu_0, \nu_0; \mu), d_\mu\phi_2(x_0, \mu_0, \nu_0; \mu))$, has a linear continuous reciprocal at $(x_0, \mu_0, \nu_0) = (\bar{x}_0, \bar{\mu}_0, \bar{\nu}_0)$.

D. The functional f has a minimum at $(\bar{x}_0, \bar{\mu}_0)$ in the class of points in $(\mathfrak{X}_0, \mathfrak{M}_0)$ which satisfy the equations

$$\phi_1[x_0, \mu_0, n(x_0, \mu_0)] = 0_u,$$

$$\psi_k(x_0, \mu_0) = 0, \qquad k = 1, \cdots, q.$$

In order to simplify the notation the element $\bar{v}$ of $\mathfrak{V}$ will be used to designate $\phi_2(\bar{x}_0, \bar{\mu}_0, \bar{\nu}_0)$; $\Phi_\alpha(x, \mu, \nu)$ to indicate $d\phi_\alpha(\bar{x}_0, \bar{\mu}_0, \bar{\nu}_0; x, \mu, \nu)$, $(\alpha = 1, 2)$; and $N(x, \mu)$ to connote $dn(\bar{x}_0, \bar{\mu}_0; x, \mu)$.

2. **The imbedding theorem.** It is of paramount importance for our purposes to show that the minimizing element $(\bar{x}_0, \bar{\mu}_0)$ is a limit point of elements (x_0, μ_0) which satisfy the equation

$$\phi_1[x_0, \mu_0, n(x_0, \mu_0)] = 0_u.$$

To show that the element $(\bar{x}_0, \bar{\mu}_0)$ is so imbedded it is convenient first to demonstrate several lemmas.

LEMMA 2.1. *In a sufficiently small neighborhood of $\bar{\mu}_0$ the equations*

(2.1)
$$\phi_1(x_0, \mu_0, \nu_0) = u,$$
$$\phi_2(x_0, \mu_0, \nu_0) = v$$

have, for each (x_0, ν_0, u, v) in a neighborhood of $(\bar{x}_0, \bar{\nu}_0, 0_u, \bar{v})$, a unique solution, $\mu_0 = M_1(x_0, \nu_0, u, v)$, of class C'.

This result follows at once from the implicit function theorems

---

* See paper IV for the definitions of terms such as composite, neighborhood, region, differential.

mentioned above,* because of assumptions C and the fact that $(\mu_0, x_0, \nu_0, u, v) = (\bar{\mu}_0, \bar{x}_0, \bar{\nu}_0, 0_u, \bar{v})$ is an initial solution of these equations.

COROLLARY 2.2. *For each (x, ν, U, V) the equations*

$$\Phi_1(x, \mu, \nu) = U,$$

$$\Phi_2(x, \mu, \nu) = V$$

have a unique continuous solution $\mu = M_2(x, \nu, U, V)$ which is equal to $dM_1(\bar{x}_0, \bar{\nu}_0, 0_u, \bar{v}; x, \nu, U, V)$.

The existence and unicity of the solution is an immediate consequence of assumption C; and the last assertion in the corollary can be established by setting $x_0 = \bar{x}_0 + ax$, $\nu_0 = \bar{\nu}_0 + a\nu$, $u = aU$, $v = \bar{v} + aV$, and $\mu_0 = M_1(x_0, \nu_0, u, v)$ in equations (2.1) and differentiating the resulting identities in a at $a = 0$.

Then to obtain the desired imbedding theorems we make a final assumption.

E. Either the function M_1, mentioned above, is independent of ν_0, or the transformation

$$\mu - d_\nu M_1[\bar{x}_0, \bar{\nu}_0, 0_u, \bar{v}; d_\mu n(\bar{x}_0, \bar{\mu}_0; \mu)]$$

has a linear continuous reciprocal.

It follows readily from Lemma 2.1 that if M_1 is independent of ν_0 and if $M(x_0, u, v)$ is defined to be $M_1(x_0, \nu_0, u, v)$, then the functions $\phi_1[x_0, M(x_0, u, v), \nu_0]$, $\phi_2[x_0, M(x_0, u, v), \nu_0]$ are independent of ν_0, at least for (x_0, ν_0, u, v) sufficiently near to $(\bar{x}_0, \bar{\nu}_0, 0_u, \bar{v})$.

THEOREM 1. *In a sufficiently small neighborhood of $\bar{\mu}_0$ the equations*

$$\phi_1[x_0, \mu_0, n(x_0, \mu_0)] = u,$$

$$\phi_2[x_0, \mu_0, n(x_0, \mu_0)] = v$$

possess a unique solution $\mu_0 = M(x_0, u, v)$ of class C' for (x_0, u, v) in some neighborhood of $(\bar{x}_0, 0_u, \bar{v})$.

From what has been said above it suffices to consider the case where M_1 is not independent of ν_0. Hence the equation

$$\mu_0 = M_1[x_0, n(x_0, \mu_0), u, v]$$

has a solution $\mu_0 = M(x_0, u, v)$ which is effective in the theorem.

Similarly a consideration of the equation

$$\mu = dM_1[\bar{x}_0, \bar{\nu}_0, 0_u, \bar{v}; x, N(x, \mu), U, V]$$

---

* See paper I, p. 65, and paper IV, p. 150.

yields the following corollary:

COROLLARY 2.3. *For each* (x, U, V) *the equations*

$$\Phi_1[x, \mu, N(x, \mu)] = U,$$

$$\Phi_2[x, \mu, N(x, \mu)] = V$$

have a unique continuous solution $\mu = M_3(x, U, V) = dM(\bar{x}_0, 0_u, \bar{v}; x, U, V)$.

COROLLARY 2.4. *Let s be an arbitrary integer,* $(x_1, \cdots, x_s)$ *a set of elements of* $\mathfrak{X}$, *and* $(\mu_1, \cdots, \mu_s)$ *a set of elements of* $\mathfrak{M}$ *satisfying* $\Phi_1[x, \mu, N(x, \mu)] = 0_u$. *Then there are functions* $x_0(b)$ *and* $\mu_0(b)$ *of class* C' *for values of* $b = (b_1, \cdots, b_s)$ *sufficiently near to* $(0, \cdots, 0)$, *which satisfy the equation*

$$\phi_1[x_0, \mu_0, n(x_0, \mu_0)] = 0_u,$$

$x_0(0) = x_0, \mu_0(0) = \bar{\mu}_0,$ *and* $\partial x_0(0)/\partial b_i = x_i, \partial \mu_0(0)/\partial b_i = \mu_i$ *for* $1 \leq i \leq s$.

Define V_i to be $\Phi_2[x_i, \mu_i, N(x_i, \mu_i)]$, $(i = 1, \cdots, s)$, $x_0(b) = x_0(b_1, \cdots, b_s) = \bar{x}_0 + b_i x_i$, and $\mu_0(b)$ to be $M[x_0(b), 0_u, \bar{v} + b_i V_i]$, where a repeated index indicates a sum. Then the first part of the corollary is obvious, and the latter part is an easy consequence of corollary 2.3.

3. The multiplier rule. The usual mode of proof utilized in the calculus of variations suffices to establish the following lemma:

LEMMA 3.1. *There are constants* $l_0, c_1, \cdots, c_q$, *not all zero, such that*

$$g(x, \mu) = l_0 df(\bar{x}_0, \bar{\mu}_0; x, \mu) + c_k d\psi_k(\bar{x}_0, \bar{\mu}_0; x, \mu) = 0$$

for every (x, μ) *satisfying the equation*

$$\Phi_1[x, \mu, N(x, \mu)] = 0_u.$$

To simplify the succeeding demonstration we shall define $g_1(x)$ and $g_2(\mu)$ to be $l_0 d_x f(\bar{x}_0, \bar{\mu}_0; x) + c_k d\text{-}\psi_k(\bar{x}_0, \bar{\mu}_0; x)$ and $l_0 d_\mu f(\bar{x}_0, \bar{\mu}_0; \mu) + c_k d_\mu \psi_k(\bar{x}_0, \bar{\mu}_0; \mu)$, respectively. It is then obvious that $g(x, \mu) = g_1(x) + g_2(\mu)$ and that g_1 and g_2 are linear continuous functionals on $\mathfrak{X}$ and $\mathfrak{M}$, respectively.

LEMMA 3.2. *The functions* $\Phi_1[x, M_3(x, U, V)]$, $N[x, M_3(x, U, V)]$, *and* $g[x, M_3(x, U, V)]$ *are independent of x and V; the latter is a linear continuous functional J on* $\mathfrak{U}$; *and for every x*

$$(3.1) \qquad\qquad g_1(x) + g_2[M_3(x, 0_u, 0_v)] = 0.$$

It follows at once from Corollary 2.3 that Φ_1 is independent of x and V. Hence we have $\Phi_1[\bar{x}, \bar{\mu}, N(\bar{x}, \bar{\mu})] = 0_u$, where $\bar{x} = x_1 - x_2$, and $\bar{\mu} = M_3(\bar{x}, 0_u, V_1 - V_2)$. Consequently we have, by Lemma 3.1, $g(\bar{x}, \bar{\mu}) = 0$ for every x_1, x_2, V_1, and V_2; this proves the rest of the lemma, since $M_3(x_1, U, V_1) - M_3(x_2, U, V_2) = M_3(x_1 - x_2, 0_u, V_1 - V_2)$.

To complete the proof of the multiplier rule we shall now prove that there is a linear continuous operation $u = P(\mu)$ on $\mathfrak{M}$ to $\mathfrak{U}$ such that $g_2(\mu) = J[P(\mu)]$ for every μ. To this end define $P(\mu)$ to be $\Phi_1[0_x, \mu, N(0_x, \mu)]$ and $Q(\mu)$ to be $\Phi_2[0_x, \mu, N(0_x, \mu)]$. Therefore, by Corollary 2.3, $\mu = M_3[0_x, P(\mu), Q(\mu)]$; as a consequence it is evident that $g_2(\mu) = g_2(M_3[0_x, P(\mu), Q(\mu)])$ which, by Lemma 3.2, is $J[P(\mu)]$.

THEOREM 2. *There exist constants $l_0, c_1, \cdots, c_q$, not all zero, and a linear, continuous, and real-valued functional J on $\mathfrak{U}$ such that*

(1) $\qquad l_0 d_\mu f(\bar{x}_0, \bar{\mu}_0; \mu) + c_k d_\mu \psi_k(\bar{x}_0, \bar{\mu}_0; \mu) - J(\Phi_1[0_x, \mu, N(0_x, \mu)]) = 0$

for every μ in $\mathfrak{M}$, and

(2) $\qquad l_0 d_x f(\bar{x}_0, \bar{\mu}_0; x) + c_k d_x \psi_k(\bar{x}_0, \bar{\mu}_0; x) - J(\Phi_1[x, 0_\mu, N(x, 0_\mu)]) = 0$

for every x in $\mathfrak{X}$.

From the previous discussion it is evident that

$$g_2(\mu) = l_0 d_\mu f(\bar{x}_0, \bar{\mu}_0; \mu) + c_k d_\mu \psi_k(\bar{x}_0, \bar{\mu}_0; \mu) = J[P(\mu)]$$

and that $P(\mu)$ is $\Phi_1[0_x, \mu, N(0_x, \mu)]$. Equation (2), the "transversality condition," follows from equation (3.1). For,

$$g_2[M_3(x, 0_u, 0_v)] = J(P[M_3(x, 0_u, 0_v)]).$$

Then making use, on the one hand, of the definitions of P and Φ_1, and on the other hand of Corollary 2.3, we obtain the desired conclusion.

4. **Instances of the theory.** We shall illustrate the generality of our theory by working out two special cases, the minima of functions of several variables and the problem of Bolza in the calculus of variations. However the theory also includes the situations treated by Hahn, Graves, and the author, as well as many others.

Consider first the problem of minimizing a real-valued function $g(x_1, \cdots, x_m)$ of m real variables in the class of points $x_0 = (x_1, \cdots, x_m)$ satisfying equations $h_k(x_0) = 0$, $(k = 1, \cdots, q)$. It is assumed that g has a minimum at a point $\bar{x}_0$ in this class and that g and h_k have first differentials at this point. Then to indicate the applicability of our theory to this case we choose the sets $\mathfrak{M}$, $\mathfrak{N}$, $\mathfrak{M}_0$, $\mathfrak{N}_0$, $\mathfrak{U}$, and $\mathfrak{V}$ to be equal to the set of real numbers, $\mathfrak{X}$ to be the m-dimensional euclidean space, and $\mathfrak{X}_0$ to be the region of definition of

the functions g and $h_1, \cdots, h_q$. We then make the following defini-
tions: $\psi_k(x_0, \mu_0) \equiv h_k(x_0)$, $(k = 1, \cdots, q)$, $f(x_0, \mu_0) \equiv g(x_0)$, $n(x_0, \mu_0) \equiv 0$,
$\phi_1(x_0, \mu_0, \nu_0) \equiv \mu_0$, and $\bar{\mu}_0 \equiv 0$. Then it is clear that $f(x_0, \mu_0)$ is a mini-
mum at $(\bar{x}_0, \bar{\mu}_0)$ in the class of all points (x_0, μ_0) in the region which
satisfy

$$\psi_k(x_0, \mu_0) = 0, \qquad\qquad k = 1, \cdots, q,$$

$$\mu_0 = 0,$$

and all previous assumptions are satisfied.

Our multiplier rule then implies the existence of constants
$l_0, c_1, \cdots, c_q$, not all zero, and a constant e such that

$$l_0 g_{x_i}(\bar{x}_0) dx_i + c_k h_{k\,x_i}(\bar{x}_0) dx_i = 0, \qquad ed\mu = 0$$

for every $dx_1, \cdots, dx_m$ and every $d\mu$; therefore we have $e = 0$ and

$$l_0 g_{x_i}(\bar{x}_0) + c_k h_{x_i}(\bar{x}_0) = 0, \qquad i = 1, \cdots, m,$$

which is the familiar form of the multiplier rule for this problem.
The problem of minimizing a function f on a region $\mathfrak{X}_0$ of a normed
linear space in the class of points satisfying functional equations
$\psi_k(x_0) = 0$, $(k = 1, \cdots, q)$, can be treated in precisely the manner
indicated above.

Let us turn our attention to the problem of Bolza in the calculus
of variations, where, for simplicity, we shall suppose that the admissi-
ble arcs are of class C'. Here we are given an arc $\bar{y}_1(x), \cdots, \bar{y}_s(x)$,
$(\bar{x}_1 \leq x \leq \bar{x}_2)$, which minimizes a functional

$$I[x_1, x_2, y, y'] = g[x_1, y(x_1), x_2, y(x_2)] + \int_{x_1}^{x_2} h[x, y(x), y'(x)] dx$$

in the class of admissible arcs which satisfy certain side conditions

$$\phi_\alpha[x, y(x), y'(x)] = 0, \qquad \alpha = 1, \cdots, r < s,$$

and end conditions

$$\psi_k[x_1, y(x_1), x_2, y(x_2)] = 0, \qquad k = 1, \cdots, q.$$

For our purposes it is sufficient to restrict attention to the
class $\mathfrak{N}_0$ of all admissible sets of functions $(y_1(x), \cdots, y_s(x))$,
$(\bar{x}_1 - e \leq x \leq \bar{x}_2 + e)$, where e is some fixed positive constant. The class
which is composed of all sets $(y_1'(x), \cdots, y_s'(x))$ of derivatives of an
element $\nu_0 = (y_1(x), \cdots, y_s(x))$ is $\mathfrak{M}_0$. Then if the transformation

$$(4.1) \qquad\qquad y_i(x) = y_i(\bar{x}_1) + \int_{\bar{x}_1}^{x} y_i'(t) dt$$

is performed upon the functionals I and ψ_k, they become functionals $f(x_0, \mu_0)$ and $\psi_k(x_0, \mu_0)$ of $x_0 = (x_1, x_2, y(\bar{x}_1))$, $\mu_0 = (y')$. Thus $\mathfrak{X}_0$ is the set of all $(x_1, x_2, y_{01}, \cdots, y_{0s})$, where $|\bar{x}_1 - x_1| < e$, $|\bar{x}_2 - x_2| < e$, and $(y_{01}, \cdots, y_{0s})$ is the value of an admissible set at $x = \bar{x}_1$.

To complete our transformation of the problem of Bolza we define $\phi_1(x_0, \mu_0, \nu_0)$ to be the set of m functions

$$\phi_1[x, \nu_0(x), \mu_0(x)], \cdots, \phi_m[x, \nu_0(x), \mu_0(x)]$$

and $n(x_0, \mu_0)$ to be the set of s functions of (x_0, μ_0) which appear in the right members of equations (4.1). It is then a simple matter to verify assumptions A through D of §1, if it is assumed that the matrix $\phi_{\alpha y_i'}$ of partial derivatives has rank m along the minimizing arc. Assumption E can be verified without difficulty by making use of the theory of Volterra integral equations.

The functional J appearing in our multiplier rule can be shown* to be representable as

$$-\int_{\bar{x}_1}^{\bar{x}_2} u_\alpha(x)\lambda_\alpha(x)dx.$$

Now, if we define F to be $l_0 h + \lambda_\alpha \phi_\alpha$, identity (1) of Theorem 2 yields the following equations:

$$F_{y_i}(x)dx + l_0(g_{y_{i1}} + g_{y_{i2}}) + c_k(\psi_{ky_{i1}} + \psi_{ky_{i2}}) = 0,$$

$$(-1)^b F(\bar{x}_b) + l_0[g_{x_b} + g_{y_{ib}}y_i'(\bar{x}_b)] + c_k[\psi_{kx_b} + \psi_{ky_{ib}}y_i'(\bar{x}_b)] = 0,$$

$$b = 1, 2;$$

moreover identity (2) becomes

$$F_{y_i'}(x) - \int_{\bar{x}_2}^x F_{y_i}(t)dt + l_0 g_{y_{i2}} + c_k \psi_{ky_{i2}} = 0.$$

By very obvious manipulations these equations can be put into the more familiar form

$$[(F - y_i' F_{y_i'})dx + F_{y_i'}dy_i]_1^2 + l_0 dg + c_k d\psi_k = 0,$$

$$F_{y_i'} = \int_{\bar{x}_1}^x F_{y_i}dx + d_i,$$

which is the multiplier rule for the problem of Bolza.

THE UNIVERSITY OF CHICAGO

* See Hahn, loc. cit., pp. 544–546.