MINIMUM PROBLEMS IN THE FUNCTIONAL CALCULUS[1]

HERMAN H. GOLDSTINE

In the year 1922 Hahn[2] published a proof of a multiplier rule for a problem more general than the well known problem of Bolza, and later, by quite different methods, Graves[3] established analogues of the multiplier rule and of the Weierstrass condition for the problem of Lagrange with integro-differential side conditions. Using Graves' method and considering a somewhat more general problem, the author[4] obtained analogues of the multiplier rule, the Clebsch condition, and the Jacobi condition. All these problems however are concerned with functionals defined over certain special function spaces. In 1934 Lusternik[5] attempted to formulate a multiplier rule for a general problem in a Banach space, but his proofs do not seem satisfactory. Two years later in a doctoral dissertation, written at the University of Kentucky, L. P. Hutchison[6] established a Lagrange rule for such a problem under natural hypotheses on the functions involved, and the author[7] proceeding in another direction generalized the statement and proof of the multiplier rule to abstract spaces.

In the present paper two problems of the functional calculus are considered in detail: the first is concerned with minimizing a functional defined on a region of a normed linear space; and the second restricts the so-called admissible points to satisfy an equation defined by a general operator. For the first of these problems it is shown that the theory of necessary and of sufficient conditions is quite like that for functions of several variables as well as for problems of the calculus of variations. In studying the second problem a multiplier rule is established under four different hypotheses, whose interrelations are studied, and the remaining necessary, as well as the sufficient conditions, are obtained under three of these hypotheses.

1. **The simple problem.** To describe the situation to be treated in

---

this section we may consider a real-valued function f which is defined on a region X_0 of a normed linear space X. The problem of this section is then that of finding a point x in the region X_0 which minimizes the function f in the class of points in X_0. Throughout this paper it will often be convenient to refer to elements in X_0 as *admissible points* and to suppose that f is of class[8] C''.

It is evident from the definitions of the first and second variations of a functional that if x is a minimizing point, then the equation

$$\delta f(x, \xi) = 0$$

must be an identity in ξ, and the lower bound of the second variation $\delta^2 f(x, \xi)$ on the unit sphere must be non-negative. This last condition is, of course, an immediate consequence of the fact that the second variation $\delta^2 f(x, \xi)$ must be non-negative.

We may now ask whether it is possible by slightly strengthening the two necessary conditions stated above to obtain a sufficiency theorem. This question is answered affirmatively in the following theorem:

THEOREM 1.1. *If x_0 is an admissible point at which the first variation $\delta f(x_0, \xi)$ vanishes identically for ξ in the space X and at which the lower bound of the second variation $\delta^2 f(x_0, \xi)$ on the unit sphere is positive, then there is a neighborhood N of the point x_0 such that $f(x) > f(x_0)$ for every $x \neq x_0$ in N.*

To prove the theorem we recall that the function f is of class C'' and hence that the second variation $\delta^2 f(x, \xi)$ is continuous in x uniformly on the unit sphere.[9] Analytically this implies that for every positive e there is a neighborhood N_e of x_0 such that for every ξ on the unit sphere and x in N_e

$$\delta^2 f(x, \xi) > \delta^2 f(x_0, \xi) - e \geq B - e,$$

where B is the lower bound of $\delta^2 f(x_0, \xi)$ for all points ξ on the unit sphere. The number B however is by hypothesis positive, and hence it is clear that the second variation will be positive at every point x in the neighborhood N_B. To complete the proof it suffices to expand $f(x)$ by means of Taylor's theorem which tells us that there is a positive number t less than unity such that

$$f(x) - f(x_0) = \delta^2 f[x_0 + t(x - x_0), (x - x_0)]/2,$$

[8] For the definitions and some properties of terms commonly used in the functional calculus see, for example, L. M. Graves, *Topics in the functional calculus*, this Bulletin, vol. 41 (1935), pp. 641–662.

[9] See Graves, loc. cit., p. 651.

since the first variation vanishes identically. This equation, combined with the result reached immediately above, then yields the conclusion of the theorem.

It is interesting to notice that in order to apply the results stated above to the calculus of variations the space X should be taken to be the set of derivatives $y'(x)$ of admissible arcs $y(x)$ and the norm of an element $y'(x)$ to be $[\int_{x_1}^{x_2}(y^2(x)+y'^2(x))dx]^{1/2}$. It can then be shown that if the strengthened form of the Legendre condition holds, our condition on the lower bound of the second variation is equivalent to the Jacobi condition, and a similar remark applies to the strengthened forms of these conditions.

2. **The more general problem.** A problem more general than the one just discussed may be stated, in terms of a function g defined on the region X_0 and having functional values in a Banach space Y, as that of finding an admissible point x which minimizes f in the class of admissible points satisfying the equation

$$(2.1) \qquad\qquad g(x) = 0_y.$$

For the purposes of the analysis it will be supposed hereafter that the space X is complete and that the function g is of class C'' on the region X_0. It will moreover be assumed that for each admissible x the contradomain, that is, the class of functional values, of $\delta g(x, \xi)$ is closed.

An admissible point x is said to satisfy the *multiplier rule* if there is a constant l and a linear, continuous, and real-valued function L defined on the space Y such that $(l, L) \neq (0, 0)$ and such that the equation

$$l\delta f(x, \xi) + L[\delta g(x, \xi)] = 0$$

is an identity in ξ; and the point x is *regular* in case the contradomain of the function $\delta g(x, \xi)$ is the entire space Y.

THEOREM 2.1. *Every admissible point x that is not regular must satisfy the multiplier rule with the constant l equal to zero.*

To prove this result we note that the contradomain of the variation $\delta g(x, \xi)$ is a proper closed subset of the space Y and hence, by a well known result concerning linear functionals,[10] there is a linear and continuous functional L defined on the space Y and not identically zero, which vanishes at each point in the contradomain of $\delta g(x, \xi)$,

[10] See, for example, S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 59.

that is, $L[\delta g(x, \xi)] = 0$. This functional is, of course, effective in the theorem, and the result is thus proved.

We may now proceed to examine in some detail the case in which x is a regular point. A preliminary result on this case is embodied in the following lemma:

LEMMA 2.1. *Let x be an admissible point which is regular. Then the first variation $\delta f(x, \xi)$ of f at the point x vanishes for every ξ satisfying the equation*

$$(2.2) \qquad\qquad \delta g(x, \xi) = 0_y,$$

if and only if the point x satisfies the multiplier rule with the constant l equal to unity and with a unique functional L.

To establish this result it is convenient to examine the operation adjoint[11] to $\delta g(x, \xi)$. This operation is defined to be the function $G(L) \equiv L[\delta g(x, \xi)]$, where L varies over the class of linear, continuous functionals on the space Y. Since the equation

$$(2.3) \qquad\qquad \delta g(x, \xi) = \eta$$

has a solution ξ for every η in Y, a theorem on the relation of a function to its adjoint[12] tells us that to every real-valued, linear, and continuous function k defined on the space X and vanishing at every point ξ which satisfies the equation (2.2) there corresponds a unique L such that $G(L) = k$. But by hypothesis δf is a functional having the same properties as k, and thus the lemma is proved, since the sufficiency of the condition in the lemma is obvious.

We now proceed to seek conditions under which the first variation $\delta f(x, \xi)$ vanishes at the solutions of the equation (2.2). To do this it is convenient to consider the following hypotheses:

(a) For every ξ satisfying (2.2) there is a function $x(b)$ of class C' defined on an interval of the real axis containing the origin, having its functional values admissible solutions of the equation (2.1), and such that $x(0) = x$, $x_b(0) = \xi$.

(b) If M is the class of all solutions of (2.2), there is a manifold N complementary to [13]M.

---

[11] Ibid., p. 99 ff.

[12] Ibid., p. 148.

[13] It is obvious that M is linear and closed; and hence, by definition, N is a linear, closed subset of X such that X is the direct sum of M and N, that is, every ξ is uniquely expressible as the sum of a μ in M and a ν in N. See F. J. Murray, *On complementary manifolds and projections in the spaces L_p and l_p*, Transactions of this Society, vol. 41 (1937), pp. 138–152.

(c) There is a linear and closed subset N of the space X in one-to-one correspondence with Y such that if η corresponds to ν then $\delta g(x, \nu) = \eta$.

(d) The space X is the composite of two Banach spaces M and N such that the partial differential $\delta_n g(x, \nu)$ has an inverse[14] at the point x.

LEMMA 2.2. *Let x be an admissible point satisfying the equation* (2.1) *and to which the hypothesis* (a) *is applicable. If then f is a minimum at x for the problem of this section, the first variation* $\delta f(x, \xi)$ *of f vanishes for every solution* ξ *of the equation* (2.2).

The proof of this lemma is obvious when one notices that the function $f[x(b)]$ must have a minimum at $b = 0$.

It is now our purpose to investigate certain relations between the hypotheses stated above. These results are contained in the following lemma:

LEMMA 2.3. *If x is an admissible solution of the equation* (2.1), *then the hypothesis* (c) *is equivalent to the statement that x is a regular point at which hypothesis* (b) *is valid, and this statement implies that the hypothesis* (a) *holds at x; in addition, hypothesis* (d) *implies* (a).

If hypothesis (c) holds at the point x, it is evident that x is regular. To prove that the linear closed set N of hypothesis (c) is effective in (b), consider an arbitrary point ξ in X, let $\eta \equiv \delta g(x, \xi)$, and let ν be the point in N corresponding to η. The point $\xi - \nu$ must then lie in the set M, and hence $\xi = \mu + \nu$. To show that this representation is unique it suffices to notice that since N is linear it contains the point 0_x and, by its definition, no other point in common with M. Conversely, it is clear that for every ν in the manifold N of hypothesis (b) there is a unique η such that $\delta g(x, \nu) = \eta$. However since x is regular, there corresponds to every η in Y a point ξ in X such that $\delta g(x, \xi) = \eta$. But ξ is uniquely expressible as $\mu + \nu$ and $\delta g(x, \mu) = 0_y$. It is therefore evident that $\delta g(x, \nu) = \eta$ and that ν is unique.

We may now show that hypothesis (c) implies (a) by considering a solution ξ of the equation (2.2). Let the point x be represented as $m_0 + n_0$ and consider the equation

$$(2.4) \qquad\qquad g(m_0 + n + b\xi) = 0_y$$

in which it is understood that n is in a neighborhood of the value n_0 and b is so near to zero that the point $m_0 + n + b\xi$ is admissible. This

---

[14] This is the assumption under which Hutchison established his multiplier rule.

equation has the initial value $(b, n) = (0, n_0)$ at which the differential with respect to n of g, $\delta g(x, \nu)$, has a continuous inverse. For, according to the hypothesis (c), the linear continuous function $\delta g(x, \nu)$ has a single-valued inverse which must be continuous since the spaces Y and N, a linear and closed subset of X, are complete.[15] There is then a unique solution $n(b)$ of (2.4) of class C' for b near zero and such that[16] $n(0) = n_0$. Differentiation of the equation (2.4) with n replaced by $n(b)$ then shows that $n_b(0) = 0_x$, since n_b is in N, and hence that the function $x(b) = m_0 + n(b) + b\xi$ is effective in hypothesis (a).

If hypothesis (d) holds, a proof very much like the one just given suffices to show that (a) also holds.

Combining Lemmas 2.1, 2.2 and 2.3, we have the following multiplier rule:

THEOREM 2.2. *Let* x *be an admissible point which is regular and satisfies the equation* (2.1). *Then if hypothesis* (a), (b), (c), *or* (d) *holds at* x *and if the function* f *is a minimum at* x *for the problem of this section, the point* x *must satisfy the multiplier rule with the constant* $l = 1$ *and with a unique functional* L.

To obtain a further necessary condition we consider a regular point x satisfying (2.1) and suppose that hypothesis (a) is applicable. If f is a minimum at x for the problem being considered, then according to Theorem 2.2 there is a unique function

$$(2.5) \qquad\qquad F(x) = f(x) + L[g(x)]$$

whose first variation $\delta F(x, \xi)$ vanishes identically, and a consideration of $F[x(b)]$, where $x(b)$ is described in hypothesis (a), serves to establish the following necessary condition:

THEOREM 2.3. *If the hypotheses of Theorem* 2.2 *hold at the point* x, *then the lower bound of the second variation* $\delta^2 F(x, \xi)$ *of the function* (2.5) *in the class of points* ξ *on the unit sphere satisfying the equation* (2.2) *is non-negative.*

3. **Sufficient conditions.** In order to establish conditions that are sufficient for a minimum it is desirable first to examine in more detail the hypotheses (c) and (d). For this purpose let x_0 be an admissible solution of the equation (2.1), which in case (c) is representable as $m_0 + n_0$ and in case (d) as (m_0, n_0). If we restrict m and n to lie near

---

[15] See Banach, op. cit., p. 41.

[16] This result follows from the Hildebrandt-Graves existence theorem for implicit functions; see Graves, loc. cit., pp. 653–661.

to m_0 and n_0, then hypothesis (d) and existence theorems concerning implicit theorems assure us that the equation

(3.1) $g(m, n) = g[(m, n)] = 0_y$

has a unique solution $n(m)$ of class C'' such that $n_0 = n(m_0)$. Consequently the point x_0 is a minimizing point for the function f in the class of admissible solutions of (2.1) if and only if the function

(3.2) $\phi(m) = f[m, n(m)]$

has a minimum at m_0 in the class of all m in a neighborhood of m_0. In the equation (3.2) immediately above, it is, of course, understood that $f(m, n)$ is the value of f at the place $x = (m, n)$. Differentiation of equations (3.2) and (3.1) yields the relations

$$\delta\phi(m_0, \mu) = \delta_m f(x_0, \mu) + \delta_n f[x_0, \delta n(m_0, \mu)],$$
$$0_y = \delta_m g(x_0, \mu) + \delta_n g[x_0, \delta n(m_0, \mu)],$$
$$\delta^2\phi(m_0, \mu) = \delta^2_{mm} f(x_0, \mu) + 2\delta^2_{mn} f[x_0; \mu, \delta n(m_0, \mu)]$$
$$+ \delta^2_{nn} f[x_0, \delta n(m_0, \mu)] + \delta_n f[x_0, \delta^2 n(m_0, \mu)],$$
$$0_y = \delta^2_{mm} g(x_0, \mu) + 2\delta^2_{mn} g[x_0; \mu, \delta n(m_0, \mu)]$$
$$+ \delta^2_{nn} g[x_0, \delta n(m_0, \mu)] + \delta_n g[x_0, \delta^2 n(m_0, \mu)].$$

If the point x_0 satisfies the multiplier rule with $l = 1$ and if $F(m, n)$ is the value of the function $F[(m, n)]$ in (2.5), then evidently $\delta_m F(x_0, \mu) = \delta_n F(x_0, \nu) = 0$ for every μ and ν; and the equations above, combined in the obvious manner, imply that

$$\delta\phi(x_0, \mu) = 0,$$
$$\delta^2\phi(x_0, \mu) = \delta^2_{mm} F(x_0, \mu) + 2\delta^2_{mn} F[x_0; \mu, \delta n(m_0, \mu)]$$
(3.3)
$$+ \delta^2_{nn} F[x_0, \delta n(m_0, \mu)]$$
$$= \delta^2 F(x_0, \xi),$$

where $\xi \equiv [\mu, \delta n(m_0, \mu)]$. By means of these relations we may now state and prove the following sufficiency theorem:

THEOREM 3.1. *Let x_0 be an admissible and regular solution of the equation (2.1) which satisfies the multiplier rule with $(l, L) = (1, L)$ and let either of the hypotheses (b), (c), or (d) hold at $x_0 = (m_0, n_0)$. Furthermore let the lower bound of the second variation $\delta^2 F(x_0, \xi)$ of the function $F(x) = f(x) + L[g(x)]$, for all ξ on the unit sphere and satisfying (2.2) be positive. Then there is a neighborhood N of the point x_0 such that $f(x) > f(x_0)$ for every $x \neq x_0$ in N that satisfies the equation (2.1).*

If hypothesis (d) is valid, then the equations (3.3) established above and the hypotheses of the theorem show that the first variation $\delta\phi(m_0, \mu)$ vanishes for all μ and that there is a positive constant B such that

$$(3.4) \qquad \delta^2\phi(m_0, \mu) = \delta^2 F(x_0, \xi) \geqq B\|\xi\|,$$

where $\xi = [\mu, \delta n(m_0, \mu)]$. Since however X is the composite of M and N, the value of $\|\xi\|$ is equal to the greater of $a\|\mu\|$ and $b\|\nu\|$, where a and b are positive constants,[17] and the relation (3.4) then implies that the hypotheses of Theorem 1.1 are satisfied at the point m_0. The theorem now follows for this case, as may be seen with the help of a remark in the first paragraph of this section.

The proof of the theorem when the hypothesis (b) or (c) holds at the point x_0 is similar to the one above but is somewhat more complicated. Since at x_0 hypothesis (c) is applicable every point ξ is uniquely expressible as $\mu(\xi) + \nu(\xi)$; and since both M and N are linear and closed, it is quite easy to see that μ and ν are linear in ξ. To show that they are also continuous we may consider a sequence ξ_n which converges to a point ξ_0 and is such that the sequence $\mu(\xi_n)$ converges to μ_0. It is then an immediate consequence of the closure of the set M and the uniqueness of the representation of each ξ that $\mu(\xi_0) = \mu_0$, and hence by a theorem on linear operators it follows that μ is continuous.[18] Similarly it may be shown that ν is continuous, and therefore since $\xi = \mu + \nu$, it is true that there is a positive constant $k \leqq 2$ such that

$$(3.5) \qquad \text{greater } [k\|\mu\|, \, k\|\nu\|] \leqq \|\xi\| \leqq \text{greater } [2\|\mu\|, \, 2\|\nu\|].$$

It now remains only to show that if we regard X as the composite of the spaces M and N and choose for the norm of (μ, ν) the greater of the numbers $a\|\mu\|$, $a\|\nu\|$, where $k \leqq a \leqq 2$, then the function g is of class C''. Having shown that g has this property, we see that the function $\phi(m)$ defined in (3.2) is well-defined and satisfies the relations (3.3) and (3.4), and therefore by the first of inequalities (3.5) that there is a positive number c such that

$$\delta^2\phi(m_0, \mu) = c\|\mu\|, \qquad\qquad \mu \text{ in } M.$$

The proof may then be completed by methods like those used before.

University of Chicago

---

[17] See Hildebrandt and Graves, *Implicit functions and their differentials in general analysis*, Transactions of this Society, vol. 29 (1927), pp. 127–153.

[18] See, Banach, op. cit., p. 41.