BULL. AUSTRAL. MATH. SOC. MOS 49B30, (49C15)

VOL. 3 (1970), 353-362.

A generalization of Lagrange multipliers

B. D. Craven

The method of Lagrange multipliers for solving a constrained stationary-value problem is generalized to allow the functions to take values in arbitrary Banach spaces (over the real field). The set of Lagrange multipliers in a finite-dimensional problem is shown to be replaced by a continuous linear mapping between the relevant Banach spaces. This theorem is applied to a calculus of variations problem, where the functional whose stationary value is sought and the constraint functional each take values in Banach spaces. Several generalizations of the Euler-Lagrange equation are obtained.

1. Constrained stationary points in a Banach space

Let $f : U \to Y$ and $h : U \to Z$ be Fréchet-differentiable maps, where X, Y, Z are Banach spaces and U is an open subset of X. Under some additional restrictions Theorem 1 gives a necessary and sufficient condition for stationarity of $f(x)$ subject to $h(x) = 0$. The proof depends on three preliminary lemmas.

LEMMA 1. *Let* S, U_0, V_0 *be real Banach spaces; let* $A : S \to U_0$ *and* $B : S \to V_0$ *be continuous linear maps, whose null spaces are* $\mathrm{N}(A)$ *respectively* $\mathrm{N}(B)$ *; let* $\mathrm{N}(A) \subset \mathrm{N}(B)$ *; let* A *map* S *onto* U_0 *. Then there exists a continuous linear map* $C : U_0 \to V_0$ *such that* $B = C \circ A$ *.*

Proof. Let p denote the projector of S onto the factor space $S/\mathrm{N}(A)$; define $A_0 : S/\mathrm{N}(A) \to U_0$ by $A_0(x+\mathrm{N}(A)) = Ax$; then A_0 is a

Received 10 August 1970.

continuous bijection of $S/N(A)$ onto U_0. So A_0^{-1} exists, continuous by Banach's bounded inverse theorem. Define similarly $B_0 : S/N(B) \to V_0$. Since $N(A) \subset N(B)$, $S/N(B)$ is a subspace of $S/N(A)$; let q denote the projector of $S/N(A)$ onto $S/N(B)$. Define $C = (B_0 \circ q) \circ A_0^{-1}$; then $C \circ A = B_0 \circ q \circ A_0^{-1} \circ A = B_0 \circ q \circ p = B$.

LEMMA 2. (Bartle [1]). *Let* X_1 *and* Z *be real Banach spaces;* S_1 *the closed ball in* X_1 *with centre* x_0*, radius* α*;* $\phi : S_1 \to Z$ *a continuously Fréchet-differentiable map, whose Fréchet derivative* $\phi'(x_0)$ *is invertible, and satisfies* $\|\phi'(x_0)\| < \frac{1}{2}\rho < \infty$. *Then there exists a constant* β *such that, if* $\|\phi(x_0)\| < \beta/\rho$*, then the equation* $\phi(x) = 0$ *has one and only one solution* $\bar{x}$ *satisfying* $\|\bar{x}-x_0\| \le \beta$.

DEFINITION 1. The linear map $M : X \to Z$, where X and Z are real Banach spaces, has *full rank* if there are subspaces X_1, X_2 of X with $X = X_1 + X_2$, $X_1 \cap X_2 = \{0\}$, $\{0\} \ne \bar{X}_1 \ne X$, such that the restriction of M to X_1 is a bijection of X_1 onto Z.
($\bar{X}_1$ = closure of X_1.)

REMARK. If X and Z have finite dimensions n, m $(m < n)$, then M has full rank iff the matrix representing M has rank m.

LEMMA 3. *Let* X, Z *be real Banach spaces;* S *an open ball in* X *with centre* 0*;* $h : S \to Z$ *a continuously Fréchet-differentiable map, for which* $h'(0)$ *has full rank, and* $h(0) = 0$. *Then to each vector* b *such that* $h'(0)b = 0$, $\|b\| = 1$ *and each sufficiently small* $\lambda > 0$, *there exists a solution* $x = \lambda b + u$ *of* $h(x) = 0$*, where* $\|u\| = o(|\lambda|)$*; and conversely every solution of* $h(x) = 0$ *for which* $\|x\|$ *is sufficiently small is of this form.*

Proof. If X is a direct sum $X_1 + X_2$, express $x \in X$ as $x = v + w$ with $v \in X_1$, $w \in X_2$. Since $h'(0)$ has full rank, $h'(0)x = Av + Bw$ where A and B are continuous linear maps and A is invertible. For fixed w, define $\phi : \bar{X}_1 \to Z$ by $\phi(v) = h(v, w)$; then $\phi'(0) = A$, which is invertible, and $\|\phi(0)\| = \|h(0, w)\| < s$ if

$\|w\| < \Delta(s) \le s$ say, since h is continuous. So by Lemma 2, for each $\varepsilon \le \beta$, $\phi(v) = 0$ has a unique solution $v = v(w)$, with $\|v-0\| < \varepsilon$, if $\|w\| < \Delta(\varepsilon/\rho)$ (where $\Delta(\varepsilon/\rho) \le \varepsilon/\rho$ may be assumed). Since h is differentiable

$$0 = h\big(v(w),\, w\big) = Av + Bw + \psi(v,\, w) ,$$

where $\|\psi(v,\, w)\| \le \varepsilon(\|v\| + \|w\|)$ if $\|v\| + \|w\| < \delta(\varepsilon)$.

Choose $\varepsilon < \frac{1}{2}\|A^{-1}\|^{-1}$ and $\varepsilon' \le \varepsilon$ such that $\varepsilon'\big(1+\rho^{-1}\big) < \delta(\varepsilon)$; if $\|w\| < \Delta(\varepsilon'/\rho)$ then $\|v\| + \|w\| < \varepsilon' + \varepsilon'/\rho < \delta(\varepsilon)$; hence

$$\|v\| = \|A^{-1}Bw+A^{-1}\psi\| \le \|A^{-1}B\|\|w\| + \|A^{-1}\|\varepsilon(\|v\|+\|w\|) ,$$

hence

$$\|v\| \le \big(\|A^{-1}B\|+\varepsilon\|A^{-1}\|\big)\|w\|/\big(1-\varepsilon\|A^{-1}\|\big) < \big(2\|A^{-1}B\|+1\big)\|w\| .$$

Therefore, taking any smaller ε and ε',

$$\|\psi\big(v(w),\, w\big)\| \le \varepsilon(\|v\|+\|w\|) < \varepsilon\big(2\|A^{-1}B\|+2\big)\|w\| = o(\|w\|) .$$

So $h(x) = 0$ has a solution

$$x = v + w = -A^{-1}Bw + w - A^{-1}\psi\big(v(w),\, w\big) = -\lambda b + u$$

where $\lambda = \|-A^{-1}Bw+w\|$, $b = \lambda^{-1}\big(-A^{-1}Bw+w\big)$, so $h'(0)b = 0$, and $\|u\| = o(|\lambda|)$; and any vector b such that $h'(0)b = 0$ is necessarily of the form $-A^{-1}Bw + w$ for some $w \in X_2$, since then $-A^{-1}Bw \in X_1$.

REMARK. If X and Z are finite-dimensional, then an application of Brouwer's fixed-point theorem proves Lemma 3 for h *differentiable* only, not necessarily *continuously differentiable*. (*Differentiable* is here taken to imply that the Fréchet derivative is a *continuous* linear mapping from X into Z.)

THEOREM 1. *Let X, Y, Z be real Banach spaces; U an open subset of X; $f : U \to Y$ a Fréchet-differentiable map, and $h : U \to Z$ a continuously Fréchet-differentiable map; assume (by restricting Y and Z) that $f(U)$ is dense in Y and $h(U)$ is dense in Z. Let $E = \{x \in U : h(x) = 0$ and $h'(x)$ has full rank$\}$. Then $f(x)$ is stationary, subject to the constraint $h(x) = 0$, at $x = a \in E$ if and* *only if there is a continuous linear map* $M : Z \to Y$ *such that*

(*) $$f'(a) = M \circ h'(a) .$$

REMARKS. $f(x)$ stationary means $f(x-\delta) - f(a) = o(\|x-a\|)$.

(*) is equivalent to the stationarity at $x = a$ of $f(x) - M \circ h(x)$ without constraints.

If $Y = \mathbb{R}$ and $Z = \mathbb{R}^m$ then M reduces to a set of m constraints, the usual Lagrange multipliers.

E is relatively open in $\{x : h(x) = 0\}$.

Proof. For $a \in E$, $f(x) - f(a) = f'(a)(x-a) + \xi$ where $\|\xi\| = o(\|x-a\|)$. By Lemma 3, $h(x) = 0$ for x in a sufficiently small neighbourhood of a if and only if $x - a = \lambda b + \eta$ where $h'(a)b = 0$, $\|b\| = 1$, and $\|\eta\| = o(|\lambda|)$; and then

$$f(x) - f(a) = \lambda f'(a)b + f'(a)\eta + \xi = \lambda f'(a)b + o(|\lambda|)$$

since $f'(a)$ is a continuous linear map. Hence, for $a \in E$,

$f(x)$ is stationary at $x = a$, subject to the constraint $h(x) = 0$

$\Leftrightarrow$ $\{h'(a)b = 0 \Rightarrow f'(a)b = 0\}$

$\Leftrightarrow$ there is a continuous linear map $M : Z \to Y$ such that $f'(a) = M \circ h'(a)$, by Lemma 1.

2. Calculus of variations in Banach spaces

Let V, S, W be (real) Banach spaces, $I = [a, b]$ a compact real interval, and $F : I \times V \times V \to S$ and $H : I \times V \times V \to W$ continuously Fréchet-differentiable maps. Let Q be a set of continuously Fréchet-differentiable functions $y : I \to V$, such that $y(b) = \beta$ and $y(a) = \alpha$ for all $y \in Q$, and such that the vector space $Q - Q$ contains $\xi(\cdot)e$ for each fixed $e \in V$ and each continuously differentiable real function ξ which vanishes on the boundary of I . Let f and h denote the maps defined, for $y \in Q$, by the Bochner integrals

$$f(y) = \int_I F(t, y(t), y'(t))dt \ ; \quad h(y) = \int_I H(t, y(t), y'(t))dt .$$

Denote by F_y and $F_{y'}$ the partial Fréchet derivatives of F with respect to its second and third arguments; for $t \in I$, $y \in Q$, denote $F_y[t, y] = F_y(t, y(t), y'(t))$ and similarly for $F_{y'}[t, y]$; denote also

$$F^+[t, y] = \int_a^t F_y[\tau, y]d\tau \; ; \; F^*[t, y] = -F^+[t, y] + F_{y'}[t, y] .$$

Denote by S_0 (respectively W_0) the closure of the range of $f'(y) : Q - Q \to S$ (respectively $h'(y) : Q - Q \to W$).

Since F is Fréchet-differentiable, so is f, and, for $y \in Q$, $\eta \in Q - Q$,

$$f'(y)\eta = \int_I \left(F_y[t, y]\eta(t)+F_{y'}[t, y]\eta'(t)\right)dt$$

$$= -\int_I F^+[t, y]\eta'(t)dt + F^+[b, y]\left(\eta(b)-\eta(a)\right) + \int_I F_{y'}[t, y]\eta'(t)dt$$

integrating by parts using Theorem 2 of [2]

$$= \int_I F^*[t, y]\eta'(t)dt + 0 .$$

LEMMA 4. *For fixed* $y \in Q$, $\int_I F^*[t, y]\eta'(t)dt = 0$ *for each* $\eta \in Q - Q \iff F^*[t, y] = 0$ *for each* $t \in I$.

Proof. Let P be the projector of S onto the one-dimensional subspace spanned by the vector $s \in S$; substitute $\eta(t) = \xi(t)e$ where $e \in V$ and $\xi(\cdot)$ is a continuously differentiable real function on I. Then, for fixed y, $P \circ F^*[t, y]\eta'(t) = \alpha(t)\xi'(t)s$, where $\alpha(\cdot)$ is a continuous function (with y as parameter). If the first statement of the lemma holds, then $\int_I \alpha(t)\xi'(t)dt = 0$ for each continuously differentiable $\xi(\cdot)$ which vanishes at a and b. By [4], page 10, Lemma 2, $\alpha(t) = 0$ for each t; therefore $P \circ F^*[t, y]e = 0$; so, since s and e are arbitrary, $F^*[t, y] = 0$. The converse is immediate.

THEOREM 2. *Let* F *and* h *be as defined above; let* E *denote the set of* $y \in Q$ *such that* $h'(y)$ *has full rank. Then* $f(y)$ *is stationary, subject to the constraint* $h(y) = 0$, *at* $a \in E$ *if and only if there is a continuous linear map* $M : W_0 \to S_0$ *such that, at* $y = a$,

$$\frac{d}{dt} K_{y'}[t, y] = K_y[t, y] \text{ , where } K = F - M \circ H \text{ .}$$

Proof. By Theorem 1, $f(y)$ is stationary, given $h(y) = 0$, at $y = a$ if and only if there is a continuous linear map $M : W_0 \to S_0$ such that $f(y) - M \circ h(y)$ has zero Fréchet derivative at $y = a$. Then (in the notation preceding Lemma 4)

$$\big(f'(y)-M \circ h'(y)\big)\eta = \int_I K^*[t, y]\eta'(t)dt \text{ .}$$

By Lemma 4, this vanishes for all $\eta \in Q - Q$ if and only if, for all $t \in I$

$$K^*[t, y] = -\int_a^t K_y[\tau, y]d\tau + K_{y'}[t, y] = 0 \text{ .}$$

If so, then $(d/dt)K_{y'}[t, y]$ exists, as a Fréchet derivative, and

$$\frac{d}{dt} K_{y'}[t, y] = K_y[t, y] \text{ .}$$

The converse is immediate.

REMARK. Theorem 2 has a partial generalization where I is replaced by a bounded closed subset of $\mathbb{R}^p$ (p-space), with boundary ∂I; and the boundary condition on $y \in Q$ becomes $y(\cdot) = \rho(\cdot)$ on ∂I, where ρ is a given function. Then $y' = \big(y_1', \ldots, y_p'\big)$ and $\eta' = \big(\eta_1', \ldots, \eta_p'\big)$ become p-vectors, mapping $\mathbb{R}^p$ into V, and F and H become functions of $t, y, y_1', \ldots, y_p'$. The proof depends on a Banach-space generalization of the Gauss-Green theorem, given in [2], Theorem A.

Let $\Phi(\cdot)$ denote the measure of $(p-1)$-dimensional surface area, used in [2]. Call a subset $E_0 \subset \mathbb{R}^p$ *thin* if $\Phi(E) < \infty$ and E_0 is a countable union of disjoint continuous images of the unit sphere in $\mathbb{R}^p$. (It follows that the p-dimensional Lebesgue measure of E_0 is zero.)

THEOREM 3. *Let f and h be as in Theorem 2, but with the compact interval I replaced by a compact subset of $\mathbb{R}^p$ whose boundary ∂I is thin; let I_0 be a thin subset of the interior of I; let Q be a set* *of continuously Fréchet-differentiable functions* $y : I \to V$ *, such that* $y(t) = \rho(t)$ *for* $t \in \partial I$ *,* $\rho(\cdot)$ *being a given function. Let* E *denote the set of* $y \in Q$ *such that* $h'(y)$ *has full rank.*

For $i = 1, 2, \ldots, p$ let the partial Fréchet derivative

$$\frac{\partial}{\partial t_i} F_{y_i'}[t, y]$$

exist at each point of $I - I_0$, and have norm integrable over I , with respect to p-dimensional Lebesgue measure. Let H satisfy similar hypotheses to F .

Then $f(y)$ is stationary, subject to the constraint $h(y) = 0$, at $y = a \in E$ if and only if there is a continuous linear map $M : W_0 \to S_0$ such that, at $y = a$,

$$\mathrm{div} K_{y'}[t, y] = K_y[t, y]$$

where $K = F - M \circ H$, $K_{y'}$ is the vector in R^p whose i-th component is $K_{y_i'}$, and

$$\mathrm{div} K_{y'}[t, y] = \sum_{i=1}^{p} \frac{\partial}{\partial t_i} K_{y_i'}[t, y] .$$

Proof. By Theorem 1, $f(y)$ is stationary, given $h(y) = 0$, at $y = a \in E$ if and only if $f(y) - M \circ h(y)$ has zero Fréchet derivative at $y = a$. Since the partial Fréchet derivatives $(\partial/\partial t_i) K_{y_i'}$ exist,

$$\sum_{i=1}^{p} K_{y_i'}[t, y]\eta_i'(t) = \mathrm{div}\{K_{y'}[t, y]\eta(t)\} - \{\mathrm{div} K_{y'}[t, y]\}\eta(t) .$$

For $y \in Q$, $\eta \in Q - Q$, and dt denoting p-dimensional Lebesgue measure,

$$\begin{aligned}(f'(y) - M \circ h'(y))\eta &= \int_I \left\{K_y[t, y]\eta(t) + \sum_{i=1}^{p} K_{y_i'}[t, y]\eta_i'(t)\right\} dt \\ &= \int_I \left(K_y[t, y] - \mathrm{div} K_{y'}[t, y]\right)\eta(t)\,dt ,\end{aligned}$$

since by [2], Theorem A,

$$\int_I \operatorname{div}\{K_{y'}[t, y]\eta(t)\}dt$$

equals an integral of $K_{y'}[t, y]\eta(t)$ over ∂I, and $\eta(t) = 0$ for $t \in \partial I$.

Since η is an arbitrary member of Q,

$$f'(y) - M \circ h'(y) = 0 \iff K_y[t, y] - \operatorname{div} K_{y'}[t, y] = 0 .$$

3. An application

Let $V = S = W = S_0 = W_0 = C(J)$, the space of all continuous complex functions on $J = [0, 1]$; let $I = [a, b]$ be a compact real interval. If $y \in Q$ maps I into V, then y is represented by a map (also denoted y) of $I \times J$ into complex numbers; define Q by requiring that $(\forall s \in J)\ y(b, s) = \beta$ and $y(a, s) = \alpha$; $(\forall s)\ y(\cdot, s)$ is continuously differentiable, and $(\forall t \in I)\ y(s, \cdot)$ is continuous. Let $P(\cdot, \cdot, \cdot,)$ be a continuously differentiable function of three real variables. For each $s \in J$, let $w(\cdot, s)$ be a (complex) measure on $I \times J$, which is weak-*-continuous in $s \in J$ and satisfies

$$\sup_{s \in J} \|w(\cdot, s)\| < \infty$$

(where norm of a measure means total variation). For $u, v \in Q$, $s \in J$, $t \in I$, define

$$F\big(t, u(t), v(t)\big)(s) = F\big(t, u(t, s), v(t, s)\big) = \int_{I \times J} P\big(t, u(t, \alpha), v(t, \beta)\big)dw\big((\alpha, \beta), s\big) .$$

Then $F\big(t, u(t), v(t)\big) \in V$, and F is continuously Fréchet-differentiable. Define H in terms of a function $\bar{P}$ and a measure $\bar{w}$, satisfying similar hypotheses to P and w.

In Theorem 2, M is a continuous linear map from $C(J)$ into $C(J)$; so by [3], Theorem 3, $K = F - M \circ H$ has the representation

$$(*)\quad K\big(t,\ y(t),\ y'(t)\big)(s) = F\big(t,\ y(t),\ y'(t)\big)(s) - \int_J H\big(t,\ y(t),\ y'(t)\big)(z)dg(z,\ s)$$

where, for each $s \in J$, $g(\cdot,\ s)$ is a (complex) measure on J, weak-*-continuous in s and satisfying

$$\sup_{s\in J} \|g(\cdot,\ s)\| < \infty .$$

Write (*) briefly as

$$K[t,\ y](s) = F[t,\ y](s) - \int_J H[t,\ y](z)dg(z,\ s) .$$

Denote by D the differential operator defined by

$$DK[t,\ y] = \frac{d}{dt} K_{y'}[t,\ y] - K_y[t,\ y] .$$

Then the criterion of Theorem 2 is formally equivalent to the following generalization of the Euler-Lagrange equation

$$(\#)\qquad DF[t,\ y](\cdot) = \int_J DH[t,\ y](z)dg(z,\ \cdot) .$$

Since, by Theorem 2, DK exists, (#) will be valid provided that also

$$\frac{d}{dt} H_{y'}[t,\ y]$$

exists (so DH exists), and (to validate differentiation under the integral sign in (#)) is, for each $t \in I$, locally dominated by a function g-integrable on I.

References

[1] Robert G. Bartle, "Newton's method in Banach spaces", *Proc. Amer. Math. Soc.* 6 (1955), 827-831.

[2] B.D. Craven, "Two properties of Bochner integrals", *Bull. Austral. Math. Soc.* 3 (1970), 363-368.

[3] B.D. Craven, "Linear mappings between topological vector spaces", (submitted for publication).

[4] I.M. Gelfand and S.V. Fomin, *Calculus of variations* (Russian; translated by R.A. Silverman). (Prentice-Hall, Englewood Cliffs, New Jersey, 1963).

[5] Kôsaku Yosida, *Functional analysis*, 2nd ed. (Die Grundlehren der mathematischen Wissenschaften, Band 123, Springer-Verlag, Berlin, Heidelberg, New York, 1968).

University of Melbourne,
Parkville, Victoria.