

A THEOREM ON CONDITIONAL EXTREMES WITH AN APPLICATION TO TOTAL DIFFERENTIALS

L. M. COURT

1. **Introduction.** In a paper [1]¹ published in 1941, the writer showed that if one function is *maximized* subject to a linear constraint, a related function is *minimized* subject to an associated linear constraint. Specifically he showed that if $\phi(q_1, \dots, q_n)$ is maximized subject to the "single" constraint $\sum_{\alpha=1}^m p_\alpha q_\alpha = m$ and $q_i = q_i(p_1, \dots, p_n)$ ($i=1, \dots, n$) are the first order conditions for this maximization, then the function $\psi(p_1, \dots, p_n) \equiv \phi[q_1(p_1, \dots, p_n), \dots, q_n(p_1, \dots, p_n)]$ is minimized subject to the same constraint except that the p 's are now regarded as the active variables. In the paper mentioned, he gave geometric and analytic proofs that the first order conditions for ψ are satisfied, but attempted only a geometric proof for the second order conditions.

In the present paper, this theorem is extended to the case in which the constraint is no longer bilinear in the two sets of variables but quite general. Moreover, we admit more than one constraint, in fact any number fewer than the number of q (or p) variables. The proof is analytic for the second order conditions as well as the first. Finally, in a supplementary section, it is shown how the two-way relationship between ϕ , ψ and the constraints sometimes can assist in the integration of total differentials.

The present theorem bears more than a passing resemblance to Mayer's law of reciprocity for isoperimetric problems in the calculus of variations [2]. There the same extremals that maximize or minimize one integral subject to a second, also maximize or minimize the second subject to the first.

2. **The theorem.** Suppose that $\phi(q_1, \dots, q_n)$ and ${}_\alpha G(q_1, \dots, q_n; p_1, \dots, p_n)$ ($\alpha=1, \dots, m; m < n$) and their first and second partial derivatives are defined and continuous for $Q \equiv (q_1, \dots, q_n) \in \mathcal{A}$ and $P \equiv (p_1, \dots, p_n) \in \mathcal{B}$, where \mathcal{A} and \mathcal{B} are domains in their respective spaces. Suppose furthermore that there is a one-to-one correspondence between the points of \mathcal{A} and \mathcal{B} ; that is, to each $Q \in \mathcal{A}$, there is a $P \equiv (p_1, \dots, p_n) \equiv (p_1^Q, \dots, p_n^Q) \in \mathcal{B}$ such that $\phi(q_1, \dots, q_n)$ is rendered stationary at Q subject to the m constraints ${}_\alpha G(q_1, \dots, q_n; p_1^Q, \dots, p_n^Q) = 0$, the p 's being regarded as parameters or inactive

Presented to the International Congress of Mathematicians, September 6, 1950; received by the editors February 6, 1950 and, in revised form, July 20, 1950.

¹ Numbers in brackets refer to the bibliography at the end of the paper.

variables in these constraints. Let $q_i = q_i(p_1, \dots, p_n)$ ($i = 1, \dots, n$) be the relations that render ϕ conditionally stationary at Q —they are the relations that express the correspondence between the points of \mathcal{A} and \mathcal{B} . Then these relations $q_i = q_i(p_1, \dots, p_n)$ —or, what amounts to the same thing, their inverses $p_K = p_K(q_1, \dots, q_n)$ ($K = 1, \dots, n$)—will also render the related function $\psi(p_1, \dots, p_n) \equiv \phi[q_1(p_1, \dots, p_n), \dots, q_n(p_1, \dots, p_n)]$ stationary at P subject to the m constraints ${}_aG(q_1^P, \dots, q_n^P; p_1, \dots, p_n) = 0$, the q 's now being regarded as the inactive variables. Moreover, if the second order conditions in Lagrange multiplier form [3] for the proper or improper maximization (minimization) of $\phi(q_1, \dots, q_n)$ subject to ${}_aG(q_1, \dots, q_n; p_1^Q, \dots, p_n^Q) = 0$ ($\alpha = 1, \dots, m$), the p 's inactive, are satisfied at Q , then the second order conditions in the same form for the proper or improper minimization (maximization) of $\psi(p_1, \dots, p_n)$ subject to the constraints ${}_aG(q_1^P, \dots, q_n^P; p_1, \dots, p_n) = 0$, the q 's inactive, will be satisfied at P .

3. **Notation.** The general q variable will always carry a lower case (small) subscript, the general p variable a capital subscript— q_i, q_t, p_K, p_S ; this convention will be all the more adhered to in summations. Differentiation with respect to a q variable will be denoted by a lower case subscript, with respect to a p variable by a capital subscript: $\phi_i = \partial\phi/\partial q_i, \psi_K = \partial\psi/\partial p_K, {}_aG_i = \partial{}_aG/\partial q_i, {}_aG_K = \partial{}_aG/\partial p_K, {}_a\lambda_i = \partial{}_a\lambda/\partial q_i, q_{iK} = \partial q_i/\partial p_K, p_{Ki} = \partial p_K/\partial q_i, \phi_{ij} = \partial^2\phi/\partial q_i\partial q_j, \psi_{KL} = \partial^2\psi/\partial p_K\partial p_L, {}_aG_{iK} = \partial^2{}_aG/\partial q_i\partial p_K$, and so forth. The umbral notation will be used, and a repeated subscript will mean summation over the natural range of the subscript—1 to m for α and 1 to n for all other subscripts.

4. **The first order conditions for ψ .** The fact that $\phi(q_1, \dots, q_n)$ is stationary at Q subject to the constraints

$$(1) \quad {}_aG(q_1, \dots, q_n; p_1, \dots, p_n) = 0 \quad (\alpha = 1, \dots, m)$$

where $P \equiv (p_1, \dots, p_n) \in \mathcal{B}$ (this is assumed) can be expressed in the Lagrange multiplier form

$$(2) \quad \phi_i - {}_a\lambda {}_aG_i = 0 \quad (i = 1, \dots, n);$$

these equations are, of course, necessarily satisfied when ϕ is a conditional extreme at Q . The second order conditions for the conditional extremization of ϕ , in Lagrange multiplier form, when ϕ is conditionally extremized (ϕ can be stationary at a point without being extremized), hinge critically upon the conditional quadratic form in the ξ 's

$$(3) \quad \{\phi_{ij} - \alpha\lambda \alpha G_{ij}\} \xi_i \xi_j$$

this quadratic form being subject to the linear constraints (again in the ξ 's)

$$(4) \quad \alpha G_i \xi_i = 0 \quad (\alpha = 1, \dots, m).$$

ϕ will be properly or improperly maximized (minimized) subject to its constraints according as this quadratic form is conditionally negative (positive) definite or semidefinite.

(1) and (2) can be regarded as equations in the "unknowns" $\alpha\lambda$ and q_i ; we know that if $\alpha\lambda^P$ and q_i^P are a particular set of values for which these equations are satisfied, then they can be solved to give the $\alpha\lambda$ and q_i as unique, single-valued functions of the p_K within some neighborhood of $P \equiv (p_1^Q, \dots, p_n^Q)$, the values of the functions (regarded as a vector) lying within some neighborhood of the augmented point $(\alpha\lambda^P, \dots, m\lambda^P; q_1^P, \dots, q_n^P)$, provided the jacobian of the left-hand members of (1) and (2) with respect to $\alpha\lambda, q_i$ does not vanish at $(\alpha\lambda^P, \dots, m\lambda^P; q_1^P, \dots, q_n^P; p_1^Q, \dots, p_n^Q)$. The jacobian in question is nothing other than the bordered determinant

$$\begin{vmatrix} 0 & \dots & 0 & {}_1G_1 & \dots & {}_1G_n \\ \cdot & \dots & \cdot & \cdot & \dots & \cdot \\ 0 & \dots & 0 & {}_mG_1 & \dots & {}_mG_n \\ \hline {}_1G_1 & \dots & {}_mG_1 & \overbrace{\phi_{ij} - \alpha\lambda \alpha G_{ij}} & & \\ \cdot & \dots & \cdot & & & \\ {}_1G_n & \dots & {}_mG_n & & & \end{vmatrix}$$

associated with the conditional quadratic form (3), (4), the jacobian and the determinant both being different from zero whenever ϕ is properly conditionally maximized or minimized. At all other points $Q \in \mathcal{A}$, the possibility of expressing the q_i (and the associated $\alpha\lambda$) in terms of the p 's is guaranteed by an explicit assumption in the hypothesis of the theorem.

By definition,

$$\psi(p_1, \dots, p_n) \equiv \phi[q_1(p_1, \dots, p_n), \dots, q_n(p_1, \dots, p_n)]$$

and $\psi_K = \phi_i q_{iK}$; also, if we think of the q 's in (1) as functions of the p 's, $\alpha G_i q_{iK} + \alpha G_K = 0$. If now we multiply (2) through by q_{iK} and use the umbral notation to sum over all i ,

$$(5) \quad \begin{aligned} 0 &= \phi_i q_{iK} - \alpha\lambda \alpha G_i q_{iK} = \psi_K - (-\alpha\lambda) \alpha G_K \\ &\equiv \psi_K - \alpha\sigma \alpha G_K \end{aligned} \quad (K = 1, \dots, n),$$

where the ${}_{\alpha}\sigma \equiv -{}_{\alpha}\lambda$ are new Lagrange multipliers. (5), of course, implies that ψ is stationary at P subject to the constraints ${}_{\alpha}G(q_1^P, \dots, q_n^P; p_1, \dots, p_n) = 0$, the q 's being regarded as inactive.

5. **The conditional extremization of ψ .** We can think of the p 's in (2) as replaced by their values in terms of the q 's—by hypothesis \mathcal{B} is mapped in a one-to-one manner onto \mathcal{A} so that not only is every q_i a single-valued function of the p 's but conversely every p_K is a single-valued function of the q 's—and then differentiate this equation with respect to q_j :

$$(6) \quad \begin{aligned} \phi_{ij} - {}_{\alpha}\lambda_j {}_{\alpha}G_i - {}_{\alpha}\lambda({}_{\alpha}G_{ij} + {}_{\alpha}G_{is} p_{sj}) &= 0 \quad \text{or} \\ \phi_{ij} - {}_{\alpha}\lambda {}_{\alpha}G_{ij} &= {}_{\alpha}\lambda_j {}_{\alpha}G_i + {}_{\alpha}\lambda {}_{\alpha}G_{is} p_{sj} \quad (i, j = 1, \dots, n). \end{aligned}$$

If the last equation is multiplied through by $\xi_i \xi_j$, summed with respect to i and j , and (4) is taken into account, we recognize the equivalence of the two quadratic forms in the ξ 's:

$$(7) \quad \{ \phi_{ij} - {}_{\alpha}\lambda {}_{\alpha}G_{ij} \} \xi_i \xi_j = {}_{\alpha}\lambda {}_{\alpha}G_{is} p_{sj} \xi_i \xi_j.$$

That is, either one of these forms, taken subject to the constraints (4), is positive or negative definite (or indefinite) if the other is.

Clearly there must be a similar equivalence between the two parallel quadratic forms in the η 's,

$$(8) \quad \{ \psi_{KL} - {}_{\alpha}\sigma {}_{\alpha}G_{KL} \} \eta_K \eta_L = {}_{\alpha}\sigma {}_{\alpha}G_{Kl} q_{lL} \eta_K \eta_L$$

subject to the parallel linear constraints

$$(9) \quad {}_{\alpha}G_K \eta_K = 0 \quad (\alpha = 1, \dots, m),$$

these linear and quadratic forms being related to ψ and the ${}_{\alpha}G$'s, regarded as functions of the active variables p_K , in the same way as the quadratic forms in (7) and the linear constraints (4) are related to ϕ and the ${}_{\alpha}G$'s, regarded as functions of the q_i .

(In the hypothesis of our theorem we have assumed the existence of only the first and second partial derivatives of ϕ and the ${}_{\alpha}G$, and some doubt may arise about whether this is sufficient to insure the existence of the ψ_{KL} in (8). Now $\psi(p_1, \dots, p_n) \equiv \phi[q_1(p_1, \dots, p_n), \dots, q_n(p_1, \dots, p_n)]$ and $\psi_{KL} = \phi_{s1} q_{sK} q_{lL} + \phi_{s2} q_{sKL}$. But the q 's are obtained as functions of the p 's by solving (1) and (2), so that at first sight the existence of the q_{sKL} , and therefore ψ_{KL} , would seem to depend on the existence of the third partials ϕ_{ijr} . However, $\psi_{KL} = {}_{\alpha}\sigma {}_{\alpha}G_{KL} + {}_{\alpha}\sigma_L {}_{\alpha}G_K + {}_{\alpha}\sigma {}_{\alpha}G_{Kl} q_{lL}$ (this equation is a strict parallel to (6) and is derived in the same way), and since ${}_{\alpha}\sigma = -{}_{\alpha}\lambda$, ψ_{KL} exists if the second partials of ϕ and the ${}_{\alpha}G$ do.)

From (1), ${}_a G_i q_{iU} = -{}_a G_U$, so that if we set

$$(10) \quad \xi_i = q_{iU} \eta_U \quad (i = 1, \dots, n),$$

$$(11) \quad {}_a G_i \xi_i \equiv {}_a G_i q_{iU} \eta_U = -{}_a G_U \eta_U = 0 \quad (\alpha = 1, \dots, m),$$

the final equation being true in virtue of (9); and since the constraints (4) are satisfied—in fact, the real meaning of (11) is that (4) and (9) are equivalent—the values assigned to the ξ_i by (10) are legitimate. If now we introduce (10) into the second form in (7) and remember that $p_{Sj} q_{jV} = \delta_{SV}$ (the Kronecker delta), we get

$$(11) \quad \begin{aligned} \alpha \lambda \, {}_a G_{iS} p_{Sj} q_{iU} \eta_U q_{jV} \eta_V &= \alpha \lambda \, {}_a G_{iS} (p_{Sj} q_{jV}) \eta_V q_{iU} \eta_U \\ &\equiv \alpha \lambda \, {}_a G_{iS} \delta_{SV} \eta_V q_{iU} \eta_U = \alpha \lambda \, {}_a G_{iV} q_{iU} \eta_V \eta_U \\ &= -\alpha \sigma \, {}_a G_{iV} q_{iU} \eta_V \eta_U, \end{aligned}$$

the final quadratic form being (because of the assumed continuity of the second partial derivatives of the ${}_a G$) the same *except for the minus sign* as the one on the right in (8). In other words, under the linear constraints (4) or their equivalent (9), either quadratic form in (8) is the negative of the corresponding quadratic form in (7), and vice-versa.

From this the second part of our theorem, having to do with the reciprocal relationship between the conditional maxima and minima of ϕ and ψ , follows immediately.

6. Application to total differentials. The result that ψ is conditionally stationary whenever ϕ is can be applied to total differentials. Let the coefficients C_i of the differential expression

$$(12) \quad C_i(q_1, \dots, q_n) dq_i \equiv [\gamma(q_1, \dots, q_n)]^{-1} \phi_i dq_i$$

satisfy $(n-1)(n-2)/2$ independent six-term integrability conditions, $\phi(q_1, \dots, q_n)$ being a parent integral. Suppose we can find m functions ${}_a G(q_1, \dots, q_n; p_1, \dots, p_n)$ such that the $m+n$ equations

$$(13) \quad {}_a G = 0, C_i - \alpha_\mu \, {}_a G_i = 0 \quad (\alpha = 1, \dots, m; i = 1, \dots, n)$$

can be solved to give the $m+n$ "unknowns" α_μ, q_i as unique single-valued functions of the p_K , the point (q_1, \dots, q_n) ranging over some domain \mathcal{A} as (p_1, \dots, p_n) ranges over a domain of its own \mathcal{B} . What this really means is that the parent integral ϕ is being rendered stationary subject to ${}_a G=0$ for $(p_1, \dots, p_n) \in \mathcal{B}$. Since $\gamma(C_i - \alpha_\mu \, {}_a G_i) \equiv \phi_i - \gamma_{\alpha\mu} \, {}_a G_i$, the $\gamma_{\alpha\mu}$ can be identified with the $\alpha\lambda$ in (2); and we see that the values of the q_i in terms of the p_K obtained by solving (13) are the same as those obtained by solving (1) and (2) if the ϕ and

the ${}_aG$ of the present section are identified with those of the earlier sections. Moreover,

$$(14) \quad \begin{aligned} C_i q_{iK} &= {}_{\alpha\mu} {}_aG_i q_{iK} = - {}_{\alpha\mu} {}_aG_K \\ &= - \gamma^{-1} {}_{\alpha\lambda} {}_aG_K = \gamma^{-1} {}_{\alpha\sigma} {}_aG_K = \gamma^{-1} \psi_K, \end{aligned}$$

so that the total differential

$$(15) \quad D_K(p_1, \dots, p_n) dp_K \equiv - {}_{\alpha\mu} {}_aG_K dp_K = \gamma^{-1} \psi_K dp_K$$

is integrable. (The D_K in (15) are defined by the identity $D_K \equiv - {}_{\alpha\mu} {}_aG_K$, in which we think of the ${}_{\alpha\mu}$ and the q_i in ${}_aG_K(q_1, \dots, q_n; p_1, \dots, p_n)$ as replaced by their values in terms of the p_K .)

The integrals $\psi(p_1, \dots, p_n)$ of $D_K dp_K$ and $\phi(q_1, \dots, q_n)$ of $C_i dq_i$ are, of course, related as the ψ and ϕ of our theorem on extremes. If (12) should happen to be difficult to integrate formally, while (15) is comparatively simple, we can first find ψ and then use the transformation between the q_i and the p_K defined by (13) to compute ϕ . (15) might be simpler to integrate than (12) because although the integrating factor is the same for both, that is, γ , the expression for γ in terms of the p_K may be simpler than the expression for γ in terms of the q_i . "Almost any" transformation can be used to introduce a new set of variables p_K in terms of the original variables q_i and then $C_i q_i dp_K$ will be integrable if $C_i dq_i$ is; the point of the present section is that if functions ${}_aG$ satisfying our assumptions can be found, the correlated total differential $D_K dp_K$ can be obtained without having to go back to $C_i dq_i$ and compute the partial derivatives of the q_i with respect to the p_K (assuming that equations (13) have been solved, we simply use the relation $D_K = - {}_{\alpha\mu} {}_aG_K$).

BIBLIOGRAPHY

1. L. M. Court, *A theorem on maxima and minima with an application to differential equations*, Journal of Mathematics and Physics vol. 20 (1941) pp. 99-106.
2. Oskar Bolza, *Lectures on the calculus of variations*, The University of Chicago Press, 1904, pp. 228, 229.
3. Harris Hancock, *Theory of maxima and minima*, Ginn, 1917, pp. 115, 116.

RUTGERS UNIVERSITY