ON THE SADDLEPOINT IN THE THEORY OF MAXIMA AND MINIMA AND IN THE CALCULUS OF VARIATIONS.

BY PROFESSOR R. G. D. RICHARDSON.

(Read before the American Mathematical Society, April 30, 1910.)

Introduction.

LAGRANGE has shown that the problem of determining a function $y(x)$ which satisfies the boundary conditions

$$(1) \qquad y(0) = y(1) = 0$$

and the integral condition

$$(2) \qquad \int_0^1 g(x, y, y')dx = 0$$

and which minimizes the integral

$$(3) \qquad \int_0^1 f(x, y, y')dx$$

is equivalent, as far as the first variation is concerned, to the problem of minimizing the integral

$$(4) \qquad \int_0^1 (f + \lambda g)dx,$$

the function being subject to no isoperimetric condition. The two constants of integration and the isoperimetric constant λ of the Lagrange differential equation

$$\frac{d}{dx}(f_{y'} + \lambda g_{y'}) + f_y + \lambda g_y = 0$$

which furnishes the solution are determined from the conditions (1) and (2). On the other hand it is possible to consider the problem of minimizing the integral (4) subject only to the boundary condition (1), in which case the minimum is obviously a function of the parameter λ. The determination of that value of λ which maximizes this minimum is a saddlepoint problem. The methods of this paper suffice to show that the first necessary condition for a solution $y(x)$ of this problem is identical with that for a solution of the foregoing isoperimetric problem.

Of a similar nature is the problem of minimizing the integral (3) for those values of $y(x)$ which satisfy the integral condition

$$\int_0^1 \{g(x, y, y') + \kappa h(x, y, y')\} dx = 0.$$

The minimum is a function $m(\kappa)$ of the parameter κ. The maximizing of $m(\kappa)$, that is, the determination of $y(x, \kappa)$ in such a way that it furnishes a maximum of a minimum is again a saddlepoint problem. It will be shown (§ 2) that the first necessary condition for a solution is identical with that for a solution of the problem of finding a minimum of (3) for those functions which satisfy the conditions

$$\int_0^1 g dx = 0, \quad \int_0^1 h dx = 0.$$

In the theory of maxima and minima there are also saddlepoint problems and related isoperimetric problems quite analogous to the foregoing and their treatment being naturally simpler is first considered in the discussion. In § 3 is given an application to an example which arises in the study of n self-adjoint linear differential equations of the second order containing n parameters.

The corresponding theory for the case of more general isoperimetric conditions is reserved for a later discussion.

§ 1. *An Extremum of an Extremum in the Theory of Maxima and Minima.*

Any solution of the problem of minimizing a function $f(x_1, \cdots, x_n)$ for those values of the variables which are subject to the relation $g(x_1, \cdots, x_n) = 0$ must satisfy the equations

$$\frac{\partial f}{\partial x_i} + \lambda \frac{\partial g}{\partial x_i} = 0 \qquad (i = 1, \cdots, n), \tag{5}$$

$$g(x_1, \cdots, x_n) = 0. \tag{6}$$

In case a solution exists, the variables $\lambda, x_1, \cdots, x_n$ are determined from the $n + 1$ equations (5) and (6).

Related to the preceding problem is the following: The minimum of the function $f(x_1, \cdots, x_n) + \lambda g(x_1, \cdots, x_n)$ is a function $m(\lambda)$ of the parameter λ; it is required to maximize this minimum considered as a function of λ. In case a minimum exists, the variables $x_1, \cdots, x_n$ are determined from the equations

$$\frac{\partial f}{\partial x_i} + \lambda \frac{\partial g}{\partial x_i} = 0 \qquad (i = 1, \cdots, n). \tag{7}$$

To determine the maximum $m(\lambda)$ it is necessary to add to (7) the condition

$$\frac{\partial}{\partial \lambda}(f + \lambda g) \equiv \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial \lambda} + \lambda \frac{\partial g}{\partial x_i} \frac{\partial x_i}{\partial \lambda} \right) + g = 0. \tag{8}$$

That the equations (7) and (8) are equivalent to (5) and (6) may be readily seen by multiplying the equations (7) by $\partial x_1/\partial \lambda, \cdots, \partial x_n/\partial \lambda$ respectively, and subtracting their sum from (8).

This result may be stated as follows: *The minimum of $f(x_1, \cdots, x_n) + \lambda g(x_1, \cdots, x_n)$, $(n > 1)$, for all values of $x_1, \cdots, x_n$ is a function of the parameter λ which when maximized gives the same constant as the minimum of f for those values of the variables which satisfy the relation $g = 0$.*

The minimum of the function $f(x_1, \cdots, x_n)$ for those values of the variables which satisfy the relations $g(x_1, \cdots, x_n) = 0$, $h(x_1, \cdots, x_n) = 0$ is found by means of the equations

$$\frac{\partial f}{\partial x_i} + \lambda \frac{\partial g}{\partial x_i} + \mu \frac{\partial h}{\partial x_i} = 0 \qquad (i = 1, \cdots, n), \tag{9}$$

$$g(x_1, \cdots, x_n) = 0, \quad h(x_1, \cdots, x_n) = 0. \tag{10}$$

The minimum of the function $f(x_1, \cdots, x_n)$ for those values of the variables which satisfy the relation

$$g(x_1, \cdots, x_n) + \kappa h(x_1, \cdots, x_n) = 0$$

is a function of the parameter κ which when maximized must satisfy the equations

$$\frac{\partial f}{\partial x_i} + \lambda \left(\frac{\partial g}{\partial x_i} + \kappa \frac{\partial h}{\partial x_i} \right) = 0, \tag{11}$$

$$g(x_1, \cdots, x_n) + \kappa h(x_1, \cdots, x_n) = 0, \tag{12}$$

$$\frac{\partial f}{\partial \kappa} \equiv \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial \kappa} = 0. \tag{13}$$

Proceeding now to show that the sets of equations (9), (10) and (11), (12), (13) are equivalent, we multiply the equations (11) by $\partial x_1/\partial\kappa, \cdots, \partial x_n/\partial\kappa$ respectively, from the sum subtract (13), and obtain

$$\lambda\sum_{i=1}^{n}\left(\frac{\partial g}{\partial x_i}\frac{\partial x_i}{\partial\kappa}+\kappa\frac{\partial h}{\partial x_i}\frac{\partial x_i}{\partial\kappa}\right)\equiv\lambda\left(\frac{dg}{d\kappa}+\kappa\frac{dh}{d\kappa}\right)=0.$$

Disregarding the case where $\lambda = 0$,* this becomes

$$\frac{dg}{d\kappa}+\kappa\frac{dh}{d\kappa}=0. \tag{14}$$

On the other hand by differentiating (12) with regard to κ, it follows that

$$\frac{dg}{d\kappa}+\kappa\frac{dh}{d\kappa}+h=0. \tag{15}$$

and from (14), (15) and (12) that

$$h(x_1, \cdots, x_n)=0, \quad g(x_1, \cdots, x_n)=0.$$

On setting $\lambda\kappa=\mu$ the complete equivalence of the two sets of equations is established and the first necessary conditions for the two problems are seen to be identical.

Since these methods of proof are perfectly general, the result may be enunciated as follows: *The minimum of the function $f(x_1, \cdots, x_n)$ for those values of the variables which satisfy the relation*

$$g(x_1, \cdots, x_n)+\sum_{i=1}^{m}\kappa_i h_i(x_1, \cdots, x_n)=0 \quad (m+1<n)$$

is a function of $\kappa_1, \cdots, \kappa_m$. The first necessary condition for a maximum of this function is identical with that for a minimum of $f(x_1, \cdots, x_n)$ for those values of $x_1, \cdots, x_n$ which satisfy the relations $g=0$, $h_1=0, \cdots, h_m=0$.

§ 2. *An Extremum of an Extremum in the Calculus of Variations.*

The problem of determining a function $y(x)$ which satisfies the boundary conditions

$$y(0)=y(1)=0 \tag{16}$$

and the integral conditions†

* In case $\lambda=0$, the solution is determined from the equations $\partial f/\partial x_i=0$ $(i=1, \cdots, n)$. The minimum is then independent of κ and the result we are seeking to establish is self-evident.

† On setting $g_1=g-c$, the case of the more general integral condition $\int_0^1 g\,dx=c$ is reduced to $\int_0^1 g_1 dx=0$.

$$(17) \qquad \int_0^1 g(x, y, y')dx = 0, \qquad \int_0^1 h(x, y, y')dx = 0$$

and which minimizes the integral

$$(18) \qquad \int_0^1 f(x, y, y')dx$$

leads to the Lagrange equation

$$(19) \qquad \frac{d}{dx}(f_{y'} + \lambda g_{y'} + \mu h_{y'}) - (f_y + \lambda g_y + \mu h_y) = 0.$$

On the other hand the minimum of the integral (18) for those functions $y(x)$ which satisfy the boundary conditions (16) and the integral condition

$$(20) \qquad \int_0^1 \{g(x, y, y') + \kappa h(x, y, y')\}dx = 0$$

is furnished by the equation

$$\frac{d}{dx}\{f_{y'} + \lambda(g_{y'} + \kappa h_{y'})\} - \{f_y + \lambda(g_y + \kappa h_y)\} = 0.$$

The condition that this minimum be maximized as a function of the parameter κ is identical with that of finding a maximum of the function

$$\phi(\epsilon_1) = \int_0^1 f(x, \bar{y} + \epsilon_1\eta, \bar{y}' + \epsilon_1\eta')dx$$

subject to the isoperimetric condition

$$\psi(\epsilon_1, \epsilon_2) = \int_0^1 \{g(x, \bar{y} + \epsilon_1\eta, \bar{y}' + \epsilon_1\eta') + (\bar{\kappa} + \epsilon_2)h(x, \bar{y} + \epsilon_1\eta, \bar{y}' + \epsilon_1\eta')\}dx = 0,$$

where $\bar{y}(x)$ and $\bar{\kappa}$ are the values of the function and parameter which furnish the extremum of the extremum, η an admissible variation arbitrarily chosen, and ϵ_1, ϵ_2 infinitesimals. On differentiating $\phi + \lambda\psi$ with regard to ϵ_1, ϵ_2 respectively and equating the results to zero for $\epsilon_1 = \epsilon_2 = 0$, one obtains the necessary conditions

$$(21) \int_0^1 [\{f_{y'} + \lambda(g_{y'} + \bar{\kappa}h_{y'})\}\eta' + \{f_y + \lambda(g_y + \bar{\kappa}h_y)\}\eta]_{y=\bar{y}}dx = 0,$$

$$(22) \qquad \int_0^1 h(x, \bar{y}, \bar{y}')dx = 0.$$

Since η is an arbitrary function vanishing at the points $x = 0$, $x = 1$, by applying the product integration usual in such cases, it is readily shown that on setting $\mu = \bar{\kappa}\lambda$, equation (21) is equivalent to (19). Since equations (20) and (22) are equivalent to (17), it follows that the first necessary condition which a solution of this saddlepoint problem must satisfy is identical with that in the ordinary isoperimetric problem with two integral conditions.

The method here employed may be applied to the other problem proposed in the introduction; it admits also of immediate generalization to functions of several variables subject to integral conditions involving several parameters.

§ 3. *Application to an Example.*

As an application of the theory of the preceding section let us consider the problem * of finding functions $u_1(x)$, $u_2(x)$ which satisfy the boundary conditions

$$(23) \qquad u_1(0) = u_1(1) = 0, \quad u_2(0) = u_2(1) = 0$$

and the integral condition

$$(24) \quad \int_0^1 [l_1(x)u_1^2(x) + l_2(x)u_2^2(x) - 1 + \kappa\,\{r_1(x)u_1^2(x) + r_2(x)u_2^2(x)\}]\,dx = 0,$$

and which minimize the integral

$$D(u_1, u_2) = \int_0^1 \{p_1(x)u_1'^2(x) + p_2(x)u_2'^2(x) - q_1(x)u_1^2(x) - q_2(x)u_2^2(x)\}\,dx,$$

where $p_i > 0$, $q_i \leqq 0$, l_i, r_i, $(i = 1, 2)$, are analytic functions of x in the interval $(0, 1)$. In order that a minimum † exist for

* Other aspects of this problem are considered by the author in a paper to be published in the *Mathematische Annalen.*

† That a minimum exists for all values of κ and is equal to the smallest (say λ_1) of the positive characteristic numbers (Eigenwerte) λ_1, λ_2, $\cdots$; λ_1', λ_2', $\cdots$ of the Lagrange equations

$$(\mathrm{A}) \qquad (p_i u_i')' + q_i u_i + \lambda\,(l_i + \kappa r_i)u_i = 0 \quad (i = 1, 2)$$

may be proved as follows: Consider the identities

$$p_i u_i'^2 - q_i u_i^2 - \lambda\,(l_i + \kappa r_i)u_i^2 \equiv \left(\frac{p_i y_i' u_i^2}{y_i}\right)' + p_i y_i^2 \left(\frac{u_i}{y_i}\right)'^2 - \frac{u_i^2}{y_i}\,[(p_i y_i')' + q_i y_i + \lambda\,(l_i + \kappa r_i)y_i] \quad (i = 1, 2),$$

various values of the parameter κ it is necessary to exclude the case that both of the functions l_1, l_2 are positive throughout the whole interval and also the case that both the functions r_1, r_2 are positive or both negative throughout the interval. (For convenience of notation in the following discussion it is assumed that r_1 takes both signs in the interval.)

If then it is possible to show that the minimum $m(\kappa)$ of $D(u_1, u_2)$ considered as a function of κ possesses a maximum, it follows from the preceding theory that the functions $u_1(x)$, $u_2(x)$ which furnish this maximum of a minimum, furnish also a minimum for the integral $D(u_1, u_2)$ subject to the boundary conditions (23) and the integral conditions

$$\int_0^1 \{l_1(x)u_1^2 + l_2(x)u_2^2\}dx = 1, \qquad \int_0^1 \{r_1(x)u_1^2 + r_2(x)u_2^2\}dx = 0.$$

In order to show that this maximum exists, we note that $D(u_1, u_2)$ can be zero only if $u_1 \equiv 0$, $u_2 \equiv 0$. Since these values will not satisfy (24), a constant $c > 0$ may be chosen such that for values of k within a restricted interval $m(k) > c$. Since m is a continuous analytic function of κ it follows that a maximum exists if it can be shown that $m(\infty) = 0$, $m(-\infty) = 0$. With this in view let us consider the function pair $u_1(x) \equiv u_1(x)$, $u_2(x) \equiv 0$. The condition (24) may then be written

---

where $u_1(x)$, $u_2(x)$ are any two functions which vanish at the end points $x = 0$, $x = 1$. If $\lambda_0 < \lambda_1$, we must show that $D(u_1, u_2) > \lambda_0$ for all functions $u_1(x)$, $u_2(x)$ which satisfy (24). The functions $y_1(x)$, $y_2(x)$ are chosen to be solutions of the differential equations

$$(p_i y_i')' + q_i y_i + \lambda_0(l_i + \kappa r_i)y_i = 0 \quad (i = 1, 2)$$

which vanish neither within the interval nor at the end points. (This is possible since λ_0 is smaller than either of the characteristic numbers λ_1, λ_1' corresponding to those solutions of the differential equations (A) which vanish at the end points. See proof by author, *Mathematische Annalen*, vol. 68, No. 2, 1910.) On setting $\lambda = \lambda_0$ in the identities, integrating and adding, we have

$$\int_0^1 \Sigma(p_i u_i'^2 - q_i u_i^2)dx - \lambda_0 \int \Sigma(l_i + \kappa r_i)u_i^2 dx = \int_0^1 \Sigma p_i y_i^2 \left(\frac{u^i}{y_i}\right)'^2 dx.$$

Since $p_1(x)$ and $p_2(x)$ are positive, this may be written $D(u_1, u_2) - \lambda_0 > 0$. That $D(u_1, u_2)$ takes the value λ_1 may be seen by setting $\lambda = \lambda_1$ in the first identity and integrating. For $u_2(x) \equiv 0$ this gives $D(u_1, u_2) - \lambda_1 = 0$ (since y_1 can be chosen equal to u_1). It may be noted that in case the characteristic numbers λ_1 and λ_1' are not equal, in order to furnish a maximum, one of the functions $u_1(x)$, $u_2(x)$ must be identically zero. In case however that $\lambda_1 = \lambda_1'$ the maximum can be furnished by the function pairs (1) $u_1(x) \not\equiv 0$, $u_2(x) \equiv 0$, (2) $u_1(x) \equiv 0$, $u_2(x) \not\equiv 0$, or (3) $u_1(x) \not\equiv 0$, $u_2(x) \not\equiv 0$.

$$(25) \qquad \int_0^1 \left\{ \frac{l_1(x)}{\kappa} + r_1(x) \right\} u_1^2(x)dx = \frac{1}{\kappa}.$$

By taking κ large enough it is possible to choose $u_1(x)$ (in an infinite number of ways) such that equation (25) is satisfied and $|u_1(x)| < \epsilon$, $|u_1'(x)| < \epsilon$, where ϵ is an arbitrarily small positive number. For this value of $u_1(x)$ (and $u_2(x) \equiv 0$), $D(u_1, u_2) < C\epsilon^2$, where C is a constant such that $C > p_1(x) - q_1(x)$. The value of κ is then so chosen that $m(\kappa)$ is less then $C\epsilon^2$. In a similar manner it may be shown that $m(-\infty) = 0$.

BROWN UNIVERSITY, PROVIDENCE,
September, 1910.

NOTE ON IDENTITIES CONNECTING CERTAIN INTEGRALS.

BY DR. LOUIS INGOLD.

(Read before the American Mathematical Society, September 7, 1910.)

BECAUSE of the general nature of the symbols and symbolic products used in the symbolic invariant theory, it is possible to apply the formulas of this theory in various special fields. In the present note the theory is employed to obtain relations connecting integrals of functions constructed out of a linearly independent set.

1. We shall be interested in functions of n parameters u_1, $u_2, \cdots, u_n$ and functions of one or more real variables $x, x_1, x_2, \cdots, y, y_1, y_2, \cdots$, etc., restricted to a definite interval, say, $0 \leq x \leq 1$. In order to distinguish readily between those quantities which are constant or functions of the parameters $u_1, \cdots, u_n$ only, and those which are functions of one or more of the variables $x, x_1, x_2, \cdots$, etc., the latter will always be denoted by black faced letters; the others by letters in ordinary type; thus $\mathbf{a}$ denotes a function of a variable x defined for values of x on the interval $0 \cdots 1$, while a denotes a function of the u_i only or a constant. The partial derivatives $\partial\mathbf{a}/\partial u_i$, $\partial a/\partial u_i$, etc., will be denoted by $\mathbf{a}_{(i)}$, $a_{(i)}$, etc.

2. Consider now the total differential of any function $\mathbf{f}(x; u_1, u_2, \cdots, u_n)$ with respect to the u_i

$$(1) \qquad d\mathbf{f} = \mathbf{f}_{(1)}du_1 + \mathbf{f}_{(2)}du_2 + \cdots + \mathbf{f}_{(n)}du_n.$$