

# GRADIENT MAPPINGS

E. H. ROTHE

1. **Introduction.** A gradient field in a finite-dimensional Euclidean space is a field of vectors  $g(x)$  for which there exists a scalar function  $I(x)$  such that  $g(x) = \text{grad } I(x)$ . A classical example is a conservative force field in which case  $I(x)$  is the potential. It has been known for a long time that the treatment of such fields offers a considerable number of simplifications and special properties not shared by arbitrary vector fields. On the other hand those boundary value and integral equation problems which can be derived from variational problems likewise offer simplifications and special properties not shared by more general problems of this type. Now the relation between an integral in the calculus of variations which is to be made an extremum and the corresponding Euler-Lagrange equation, or rather the operator given by the Euler-Lagrange expression, is quite analogous to the relation between the scalar  $I(x)$  and its gradient. It seems therefore reasonable to expect that with a proper definition of the term gradient one will be able to obtain a theory which encompasses the finite-dimensional as well as the function space case.<sup>1</sup>

2. **The definition of the term gradient mapping.** If  $I(x)$  is a differentiable scalar defined for points  $x = (x_1, x_2, \dots, x_n)$  of a Euclidean  $n$ -space  $E$ , then the differential  $dI$  corresponding to the increment  $h = (h_1, h_2, \dots, h_n)$  is given by

$$(2.1) \quad dI = \sum_{\nu=1}^n \partial I / \partial x_\nu h_\nu = (\text{grad } I, h)$$

where the parentheses indicate the scalar product. Thus  $g(x) = \text{grad } I(x)$  assigns to the point  $x$  of  $E$  the linear form  $l_x(h) = dI = (g(x), h)$ , that is, an element of the conjugate space  $E^*$  of  $E$ . This remark motivates the following definition.

**DEFINITION 2.1.** Let  $E$  be a real Banach space.<sup>2</sup> Let  $I = I(x)$  be a scalar function defined in  $E$  or part of  $E$ . Suppose that  $I$  possesses a continuous Fréchet differential, that is, that there exists a linear

---

An address delivered before the Norman meeting of the Society on November 23, 1951 by invitation of the Committee to Select Hour Speakers for Western Sectional Meetings; received by the editors December 24, 1951.

<sup>1</sup> The use of gradients in Hilbert space for the treatment of functional equations seems to occur first in [9]. (Numbers in brackets refer to the bibliography at the end of the paper.)

<sup>2</sup> For the definition of a Banach space see, for example, [14, p. 10].

bounded functional of  $h$ ,  $l_x(h) = D(x, h)$ , depending continuously on  $x$  such that

$$(2.2) \quad I(x + h) - I(x) = D(x, h) + R(x, h)$$

where with the usual notation  $\|h\|$  for the norm of  $h$

$$(2.2') \quad \lim_{\|h\| \rightarrow 0} R(x, h)/\|h\| = 0.$$

If  $E^*$  denotes the conjugate space [14, p. 21] of  $E$ , then the Fréchet differential  $D(x, h) = l_x(h)$  induces a mapping of  $E$  (or part of  $E$ ) into  $E^*$  which maps the point  $x$  of  $E$  into the point  $l_x(h) = l_x$  of  $E^*$ . Any mapping of  $E$  (or part of  $E$ ) into  $E^*$  which is thus induced by the Fréchet differential of a scalar is called a gradient mapping.

If  $l_x = l_x(h)$  is a given map of  $E$  (or part of it) into  $E^*$  which has a continuous Fréchet differential (with respect to  $x$ )  $\delta_x(h, k)$ , then the symmetry of the bilinear form (in  $h$  and  $k$ )  $\delta_x(h, k)$  is a necessary and sufficient condition that  $l_x$  is a gradient map. This follows immediately from a theorem by Kerner [15, Satz 2'].

If  $E$  is especially a Hilbert space,<sup>3</sup> it is well known that to a given linear bounded functional  $l(h)$  there exists a uniquely determined element  $g$  of  $E$  such that  $l(h) = (g, h)$ , and that one may consider  $E$  as identical with  $E^*$  by identifying  $l$  with  $g$ . In this sense we consider in a Hilbert space as the gradient of  $I(x)$  the element  $g(x)$  of  $E$  which is uniquely determined by

$$D(x, h) = l_x(h) = (g(x), h).$$

Thus  $dI = (g(x), h)$  in obvious generalization of (2.1).

If  $G(x)$  is a given map of the Hilbert space  $E$  or part of it into itself which has a continuous Fréchet differential (with respect to  $x$ )  $L_x(k)$ , then a necessary and sufficient condition that  $G(x)$  is a gradient is that the bounded linear operator (in  $k$ )  $L_x(k)$  be symmetric, that is, that  $(L_x(k), h) = (k, L_x(h))$ . This is a nearly immediate consequence<sup>4</sup> of the criterion for general Banach spaces stated above. In the special case that the given  $G(x)$  is a bounded linear operator, its Fréchet differential  $L_x(k) = G(k)$ . Thus the necessary and sufficient condition for a bounded linear operator to be a gradient is its symmetry.<sup>5</sup> From this point of view gradients appear as a natural generaliza-

<sup>3</sup> A Hilbert space is a Banach space in which for any two elements  $x, y$  a scalar  $(x, y)$  satisfying the usual rules is defined such that  $(x, x)^{1/2}$  is the norm of  $x$ .

<sup>4</sup> See [26, Lemma 2.3]; compare also [22, Lemma 3.1 and the "Remark to Lemma 3.1," p. 585], where slightly different definitions were used.

<sup>5</sup> For the treatment of eigenvalue and boundary value problems connected with such operators, we refer to the classical treatment in [7].

tion of symmetric linear operators.<sup>6</sup>

**3. Examples.** We list a few examples of scalars and their gradients. To most of these examples we shall have to refer later on.

**EXAMPLE 1.** Let  $A(x)$  be a linear bounded symmetric operator defined in a Hilbert space  $E$ , and let

$$(3.1) \quad I(x) = (1/2)(A(x), x).$$

Then  $\text{grad } I = A(x)$ . Note that the remainder term  $R(x, h)$  of (2.2) becomes in our case

$$(3.2) \quad R(x, h) = (A(h), h)/2$$

and therefore due to the supposed boundedness of  $A$  certainly satisfies (2.2'). As a special case we mention

$$(3.3) \quad I(x) = \frac{1}{2} \int_D \int_D K(s, t) x(s) x(t) ds dt,$$

$$(3.4) \quad \text{grad } I = A(x) = \int_D K(s, t) x(t) dt \quad (K(s, t) = K(t, s)),$$

where  $D$  is a bounded domain in a Euclidean  $n$ -space,  $s = (s_1, s_2, \dots, s_n)$  a point of  $D$ , and the integration is extended over  $D$ . The Hilbert space  $E$  is the space of all functions  $x(s)$  square integrable over  $D$ . Under certain assumptions about the kernel  $K(s, t)$  the symmetric operator (3.4) is indeed bounded.<sup>7</sup> Under somewhat more restrictive conditions, e.g., if

$$(3.5) \quad \int_D \int_D K(s, t)^2 ds dt < \infty,$$

the operator is completely continuous.<sup>8</sup>

<sup>6</sup> Thus, the operators  $G(x)$  in a Hilbert space called by Lusternik [17] symmetric and homogeneous of order  $p-1$  are in the terminology of the present paper gradients for which the corresponding scalars  $I(x)$  are homogeneous, that is, satisfy equation (4.5) of the present paper. In generalization of the classical case of linear positive definite symmetric operators ( $p=2$ ), the existence of infinitely many positive eigenvalues  $\lambda_i$  with  $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots$  and  $\lim_{i \rightarrow \infty} \lambda_i = 0$  is proved in [17] for operators  $G$  which are homogeneous with even  $p$ , completely continuous and positive, that is, for which  $(G(x), x) > 0$  for  $x \neq 0$ . Soboleff [32] obtains a similar result for symmetric, completely continuous, positive definite Hilbert space operators  $G(x)$  with  $G(-x) = -G(x)$ .

<sup>7</sup> See, for example, [30, pp. 32 and 36].

<sup>8</sup> A map of a Banach space  $E$  (or part of it) into the Banach space  $E_1$  (which may coincide with  $E$ ) is called completely continuous if (i) it is continuous, and (ii) the image of every bounded set is compact in  $E_1$  (in the sense of the definition given in [2, p. 84]).

EXAMPLE 1a. With the same notations as in example 1, let  $A(x)$  be an unbounded linear symmetric operator defined in a linear subset of the Hilbert space  $E$ . In this case we can no longer conclude that the remainder term (3.2) satisfies (2.2'), and therefore  $A(x)$  is in general not the gradient of the scalar (3.1). Consider, for example, the linear symmetric operator

$$(3.6) \quad A(y) = B(y) + y$$

with

$$(3.7) \quad B(y) = - \sum_{\nu=1}^n \frac{\partial}{\partial s_\nu} \left( p \frac{\partial y}{\partial s_\nu} \right).$$

Here  $y=y(s)$  is a function of the point  $s=(s_1, s_2, \dots, s_n)$  of the domain  $D$  which vanishes on the boundary of  $D$  and for which the derivatives occurring in  $B$  exist and are continuous while  $p=p(s)$  is a given positive differentiable function of  $s$ . Application of Green's formula shows that the scalar (3.1) becomes the generalized Dirichlet integral

$$(3.8) \quad \frac{1}{2} (A(y), y) = \frac{1}{2} \int_D \left\{ p \sum_{\nu=1}^n (\partial y / \partial s_\nu)^2 + y^2 \right\} ds.$$

EXAMPLE 2. Let  $E$  be the space of all functions  $x=x(t)$  which together with their first derivative are defined and continuous in  $0 \leq t \leq 1$ . With the norm

$$\|x\| = \max |x(t)| + \max |x'(t)|$$

$E$  becomes a Banach space [4, pp. 11 and 54]. Let

$$I(x) = \int_0^1 f(t, x(t), x'(t)) dt.$$

Then under proper assumptions about the given function  $f$  the Fréchet differential  $dI$  corresponding to the increment  $h=h(t)$  is

$$\begin{aligned} l_x(h) &= \int_0^1 \{ h(t) f_x(t, x(t), x'(t)) + h'(t) f_{x'}(t, x(t), x'(t)) \} dt \\ &= h(t) f_{x'}(t, x(t), x'(t)) \Big|_{t=0}^{t=1} + \int_0^1 h(t) \left\{ - \frac{d}{dx} f_{x'} + f_x \right\} dt. \end{aligned}$$

This is a linear functional on  $E$ , i.e., an element of  $E^*$  and the gradient mapping  $g(x)$  is the map which maps  $x=x(t)$  into the element  $l_x(h)$  of  $E^*$ .

**EXAMPLE 3.** The Hammerstein scalar. Let  $D$ ,  $x(s)$ , and  $K(s, t)$  have the same meaning as in example 1. Moreover we suppose that  $K(s, t)$  satisfies (3.5), is symmetric and positive definite. Finally let  $f(t, u)$  be a continuous function defined for  $t \in D$  and for  $-\infty < u < \infty$ . The Hammerstein scalar  $i = i(x)$  is then defined by

$$(3.9) \quad i(x) = \frac{1}{2} \int_D \int_D K(s, t) x(s) x(t) ds dt + I(x)$$

where

$$(3.10) \quad I(x) = \int_D \left\{ \int_0^{v(t)} f(t, u) du \right\} dt$$

with

$$(3.11) \quad y(t) = \int_D K(s, t) x(s) ds.$$

The significance of this scalar lies in its connection with the so-called nonlinear integral equation of the Hammerstein type

$$(3.12) \quad y(s) + \int_D K(s, t) f(t, y(t)) dt = 0.$$

It can be shown<sup>9</sup> that  $\text{grad } i(x)$  equals the left member of (3.12) (with  $y$  expressed in terms of  $x$  by means of (3.11)) if in addition to the above assumptions concerning  $K$  and  $f$  at least one of the following conditions (A), (B) is satisfied:

- (A)  $|K(s, t)|$  is bounded;
- (B)  $|f|$  is dominated by a linear function of  $u$  with coefficients independent of  $t$ .

For later reference we note that if  $f$  satisfies (B) and has in addition a continuous partial derivative with respect to  $u$ , and if  $d^2(x, h, h_1)$  denotes the second Fréchet differential of  $i(x)$  corresponding to the increments  $h = h(t)$  and  $h_1 = h_1(t)$ , then

$$(3.13) \quad \begin{aligned} d^2(x, h, h) &= \frac{1}{2} \int_D \int_D K(s, t) h(s) h(t) ds dt \\ &+ \int_D \int_D \partial f / \partial u(t, y(t)) \left\{ \int_D K(s, t) h(s) ds \right\}^2 dt, \end{aligned}$$

where again  $y(t)$  is given by (3.11) [26, equation (7.24)].

<sup>9</sup> [25, Theorem 3.1]. See [25] also for a more precise definition of the Hammerstein scalar; in the terminology of §7 of the present paper the Hammerstein scalar should be defined in the extension  $E_1$  of  $E$ .

**4. The existence of an extremum.** The examples of the previous section will make clear the role of gradients in the treatment of boundary value problems or integral equations: if the scalar  $i(x)$  has an extremum in an interior point of the solid sphere  $V = V_R$  given by  $\|x\| \leq R$  of the Banach space  $E$ , then

$$(4.1) \quad \text{grad } i = g(x) = 0;$$

if  $x$  is an extremum point on the boundary  $S = S_R$  of  $V_R$ , then it is easy to see (at least in the case that  $E$  is a Hilbert space) that for some  $\lambda$

$$(4.2) \quad g(x) + \lambda x = 0.$$

The "variational" method of finding solutions of (4.1) and (4.2) consists in finding extrema of  $i(x)$ . In a finite-dimensional space  $i(x)$  reaches a maximum and minimum in every bounded closed set if  $i(x)$  is continuous, and at least a minimum if  $i(x)$  is lower semicontinuous. As is well known, these statements are no more true in function spaces. If, for example,  $E$  is the Hilbert space of points  $x = (x_1, x_2, \dots)$  with  $\sum_{i=1}^{\infty} x_i^2 < \infty$ , and  $\lambda_1, \lambda_2, \dots$  a sequence of positive nonincreasing constants converging to zero, then the scalar

$$(4.3) \quad I(x) = \sum_{p=1}^{\infty} \lambda_p x_p^2$$

is continuous and positive on the unit sphere  $S_1$  given by  $\|x\| = 1$ . On the other hand the greatest lower bound of  $I(x)$  on  $S_1$  is 0, as is seen easily by considering the sequence of points  $x^n$  on  $S_1$  whose coordinates are all 0 except for the  $n$ th, which is 1. Thus  $I(x)$  has no minimum on  $S_1$ .

This situation is due to the fact that a closed bounded set in  $E$  is not necessarily compact. However, it follows from a well known theorem of Alaoglu [1] that the solid sphere  $V$  is compact in the "weak topology" if  $E$  is reflexive,<sup>10</sup> where the weak topology is the one induced by the following neighborhood definition: let  $K$  be a number greater than  $R$ ; then a neighborhood of a point  $x_0$  of  $V_K$  is determined by a positive number  $\epsilon$  and a finite number of continuous linear functionals  $l_i(x)$  ( $i = 1, 2, \dots, n$ ) and consists of all points  $x$  of  $V_K$  for which

$$|l_i(x) - l_i(x_0)| < \epsilon \text{ for } i = 1, 2, \dots, n.$$

---

<sup>10</sup> For the definition of a reflexive Banach space, see [14, p. 22]. Any Hilbert space is reflexive. More generally any uniformly convex space is reflexive; see [19].

As a consequence the scalar  $I(x)$  takes a maximum and a minimum in  $V_R$  if it is weakly continuous (that is, continuous in the weak topology) and it takes a minimum in  $V_R$  if it is weakly lower semicontinuous. It follows that for any integer  $n$

$$(4.4) \quad i(x) = \|x\|^n + I(x),$$

with weakly continuous  $I(x)$ , takes a minimum in  $V$  since  $\|x\|^n$  is weakly lower semicontinuous [24, §2].

If  $I(x)$  in addition to being weakly continuous is homogeneous, that is, if for some non-negative  $p$

$$(4.5) \quad I(\alpha x) = \alpha^p I(x),$$

statements about the existence of extrema on the boundary  $S_R$  of  $V$  can be made. If, for example,  $I(x)$  takes positive values at all in  $V_R$ , it takes a positive maximum in  $V_R$ . It then follows easily from (4.5) that this positive maximum must be taken on  $S_R$ , and consideration of  $-I(x)$  shows that in any case an extremum is taken on  $S_R$ . Thus the scalar (4.3) takes the maximum  $\lambda_1$  in the point

$$(1, 0, 0, \dots)$$

of  $S_1$ , but no minimum on  $S_1$ . It takes its minimum 0 in  $V_1$  in the center 0 of this solid sphere.

**5. Weak continuity of  $I(x)$  and complete continuity of  $\text{grad } I(x)$ .** The next task is to find sufficient conditions for the weak continuity of  $I(x)$ . In this respect the following theorem holds in any Banach space  $E$  [24, Theorem 3.2]: sufficient for the weak continuity of  $I(x)$  is the complete continuity of the gradient map  $G(x)$  (that is, that  $G(x)$  is continuous, and that the image of  $V$  under  $G$  is a compact set of  $E$ ; compare the definition of a gradient map in §2). More explicitly: if  $G(x)$  is completely continuous, then there exists to a given positive  $\eta$  linear continuous functionals  $l_1, l_2, \dots, l_n$  and an  $\epsilon > 0$  such that  $|I(x+h) - I(x)| < \eta$  for all  $h$  with  $x+h \in V_K$  for which  $|l_i(h)| < \epsilon$  ( $i=1, 2, \dots, n$ ) where  $V_K$  has the same meaning as in §4. Under the same assumption for complete continuity of  $G$  it is also true that to given  $\eta > 0$  there exist linear continuous functionals  $l_i$  ( $i=1, \dots, n$ ) such that

$$|I(x+h) - I(x)| < \eta \|h\|/2 \text{ if } \|l_i(h)\| < \eta \|h\|/2 \quad (i = 1, 2, \dots, n).$$

If  $E$  is a Hilbert space, the converse of the last statement holds [24, Theorem 3.3].

As a special application we draw the conclusion that the scalar (3.1) is weakly continuous if the operator  $A(x)$  is completely con-

tinuous.<sup>11</sup> Since this scalar is homogeneous (the  $p$  of (4.5) being 2), it follows from the previous section that this scalar reaches an extremum on the surface of the unit sphere. This is especially true of the double integral (3.4) (under the assumption (3.5)), a well known fact which is basic for the variational treatment of the theory of linear integral equations.<sup>12</sup>

**6. Extension of the domain of applicability of the preceding theory.** The application of the preceding theory seems to be confined to those scalars which have a completely continuous gradient, and therefore to have no bearing, for example, on problems connected with the differential operator  $A$  defined by (3.6); for (3.7) is not completely continuous quite apart from the fact already pointed out in example 1a that this operator is not the gradient of the scalar (3.8) in the sense of the definition given in §2.

However, it is sometimes possible by introducing a new "independent variable" and a new norm<sup>13</sup> to make the gradient completely continuous and therefore the corresponding scalar gradient weakly continuous. We illustrate this procedure for the operator  $A$  defined by (3.6). It is well known that, under certain conditions concerning the differentiability of  $p$  and of the boundary of the bounded domain  $D$ , there exists a Green's function  $K(s, t)$  belonging to the operator  $B$  defined in (3.7) and the boundary conditions zero.<sup>14</sup> For every continuously differentiable function the unique solution of the boundary value problem

---

<sup>11</sup> That the complete continuity of the linear operator  $A(x)$  is a necessary and sufficient condition for the weak continuity of the scalar (3.1) was already recognized by F. Riesz [20, pp. 96–97, footnote]. As regards the terminology, it should be noted that Riesz (following Hilbert) calls the quadratic form  $(A(x), x)$  completely continuous if it is weakly continuous in the sense explained in §4 of the present paper. The terminology which calls a nonlinear (not necessarily quadratic) scalar in Hilbert space completely rather than weakly continuous if its gradient is completely continuous is still used in [22, Definition 5.1 and Theorem 5.1] and in [23]. A scalar which is completely continuous in the sense of this latter definition is actually weakly continuous (as noticed by the referee of [23]; compare Lemma 2.1 of that paper). The connection between weak continuity of a scalar and the complete continuity of its gradient has also been noticed by E. S. Citlanadze in the Hilbert space case [5, Theorem 2] and in the case of a reflexive Banach space with basis in [6].

<sup>12</sup> See, for example, [7, vol. 1, chap. 3].

<sup>13</sup> The method of renorming a Hilbert space  $E$  (or subspace of  $E$ ) was introduced and applied to the treatment of second order boundary value problems by K. Friedrichs [8]. For higher order problems, compare [3, §10]. For "introducing a new variable," see also [12, p. 68].

<sup>14</sup> Schauder [29, p. 279] showed that the use of the Green's function is not necessary for the method in question.

$$(6.1) \quad B(y) = x, \quad y = 0 \text{ on the boundary of } D,$$

is then given by (3.11) for which we write shortly

$$(6.2) \quad y = K(x).$$

Moreover, if  $E$  is the Hilbert space of all functions  $x(t)$  for which

$$(6.3) \quad \int_D x^2(t) dt$$

exists with the square root of this integral as norm, then (6.2) represents a completely continuous linear map of  $E$  into itself mapping  $x$  into  $y$ .

We now introduce  $x$  as a "new variable" instead of  $y$  in the operator  $A$  and obtain from (3.6), (6.1), and (6.2) (at least for a certain subspace of  $E$ )

$$(6.4) \quad A(y) = x + K(x).$$

We now renorm our space as follows: with the scalar product corresponding to the norm given by (6.3), we have from (6.1), (6.2), (3.7), and Green's formula

$$(x, K(x)) = (B(y), y) = \int_D p \sum_{\nu=1}^n (\partial y / \partial s_\nu)^2 ds > 0.$$

We therefore can define a norm  $\|x\|$  given by

$$(6.5) \quad \|x\|^2 = (x, K(x))$$

instead of (6.3).  $K(x)$  is defined for all  $x$  in  $E$ . However,  $E$  is not necessarily complete in the new norm given by (6.5). But  $E$  can be completed.<sup>15</sup> We denote the completed space with the new norm by  $E_1$ . If we denote the scalar product corresponding to the new norm  $\|x\|$  by square brackets such that  $\|x\|^2 = [x, x]$ , we obtain from (6.4) and (6.2) for the scalar (3.8), which we now denote by  $i(x)$ ,

$$(6.6) \quad \begin{aligned} 2i(y) &= (A(y), y) = (x, K(x)) + (K(x), K(x)) \\ &= \|x\|^2 + [x, K(x)]. \end{aligned}$$

Now  $K(x)$  is completely continuous in the norm given by (6.3), and it can be seen easily that it remains completely continuous in the new norm  $\|x\|$ . Consequently, the scalar  $I(x)$  defined by

$$(6.7) \quad I(x) = [x, K(x)]/2$$

<sup>15</sup> Concerning the completion of  $E$ , see [8]. For a more concrete method applicable if the eigenfunctions of the positive definite kernel are known, see [25].

is weakly continuous by the last paragraph of §5, and the scalar (3.8) is by (6.6) of the form (4.4) that is,

$$(6.8) \quad i(x) = \|x\|^2 + I(x)$$

with weakly continuous  $I$ , and it takes, therefore, a minimum in every sphere  $\|x\| \leq R$  while the scalar (6.7) takes not only a minimum in such a sphere but also a maximum on its boundary  $\|x\| = R$  (compare the last paragraph of §4).<sup>16</sup>

**7. Application to the Hammerstein integral equation.** The method illustrated in the previous section in the case of the linear operator  $A$  can also be applied to nonlinear problems. We shall illustrate this for the nonlinear integral equation (3.12) with the symmetric positive definite kernel  $K(s, t)$  satisfying the assumptions made in §3, example 3. We define the spaces  $E$  and  $E_1$  as in §6 with  $K$  now being the kernel of (3.12).<sup>17</sup> From (6.5) the Hammerstein scalar  $i(x)$  defined by (3.9), (3.10) can then be written

$$(7.1) \quad i(x) = \frac{\|x\|^2}{2} + I(x),$$

where  $I(x)$  is defined by (3.10), (3.11) with continuous  $f$ . Then, as has been mentioned already in §3, either of the conditions (A), (B) (§3, example 3) is sufficient for the left member of (3.12) (expressed by means of (3.11) in terms of  $x$ ) to be equal to  $\text{grad } i(x)$ . It can also be proved that either of these conditions is sufficient for

$$\text{grad } I(x) = \int_D K(s, t) f(t, y(t)) dt$$

to be completely continuous [25, Theorem 3.2]. Thus, the Hammerstein scalar (7.1) has the same properties as (4.4) and therefore reaches a minimum in any solid sphere  $V_R$  of  $E_1$ . If the minimum is taken in an interior point  $x_0$  of  $V_R$ , the  $\text{grad } i$  will be zero in  $x_0$ , that is, the  $y$  connected with  $x = x_0$  by (3.11) will be a solution of (3.12). The minimum point  $x_0$  will certainly, at least for large enough  $R$ , be an interior point of  $V_R$  if  $i(x) \rightarrow \infty$  as  $\|x\| \rightarrow \infty$ . Now it can be shown that this latter condition is satisfied if

<sup>16</sup> The point (or points)  $x$  in which the extremum is attained is in  $E_1$  but not necessarily in  $E$ . For the applications it is therefore of importance to extend the operator (6.2) from  $E$  to  $E_1$  in such a way that  $y$  is in the original function space. Compare the following footnote.

<sup>17</sup> For a more precise description of the linear operator (6.2) (that is, (3.11)) in the extended space  $E_1$ , see [25, §§2 and 3].

$$(7.2) \quad \int_0^y f(t, u) du > -k/2y^2 - C,$$

where  $C$  and  $k$  are constants, the latter subject to the condition  $0 \leq k < 1/\lambda_1$  with  $\lambda_1$  being the greatest eigenvalue of the kernel  $K$ . Thus, the conditions of this section taken together (with either  $A$  or  $B$ ) are sufficient for the existence of a solution of the Hammerstein equation (3.12).<sup>18</sup>

**8. Critical points and Leray-Schauder index.** So far the existence theorems for boundary value or integral equation problems which can be written in form (4.1) have been based on the proof that the corresponding scalar  $i(x)$  has a maximum or minimum. However, maxima or minima are rather special types of critical points, i.e., of points  $x$  satisfying (4.1). Let  $x_0$  be a point which is not a limit of critical points. Then a sufficient condition for  $x_0$  to be a critical point is that at least one of the (Morse)-type numbers at  $x_0$  are different from zero. These numbers are defined as follows:<sup>19</sup> let  $U$  be a neighborhood of  $x_0$  containing no critical point except possibly  $x_0$ . Let  $C$  be the intersection of  $U$  with the set of those  $x$  for which  $i(x) \leq i(x_0)$ , and  $C'$  the set  $C$  minus the point  $x_0$ ; the  $r$ th type number  $m^r$  of  $i(x)$  at  $x_0$  is then defined as the singular  $r$ th relative Betti number of  $C \bmod C'$  ( $r=0, 1, \dots$ ). In what follows we shall always assume that  $E$  is a Hilbert space, and that  $g(x) = \text{grad } i(x)$  is of the form

$$(8.1) \quad g(x) = x/2 + G(x)$$

with completely continuous  $G$ . By the result mentioned at the beginning of §5,  $i(x)$  is then of the form (4.4) with  $n=2$  and weakly continuous  $I(x)$  encountered several times in our examples. Now for mappings of the form (8.1) the mapping degree is defined. More precisely: if  $D$  is any open bounded convex domain in  $E$ , and if  $y$  is a point of  $E$  which does not lie on the boundary of the image of  $D$  under  $g$ , then the degree  $j = j(D, g, y)$  of the map  $g: D \rightarrow E$  at  $y$  is defined [16]. Therefore, especially  $j = j(g, U, 0)$  is defined where  $0$  denotes the zero point of  $E$ , and  $U$  is any neighborhood of the point

<sup>18</sup> This is one of the principal results of Hammerstein's paper [13]. Since then more general results concerning existence and uniqueness of solutions of equation (3.12) and systems of such equations have been obtained by M. Golomb in [9] and [10] without the assumption of positive definiteness and of symmetry. See also [21, §5]. The method described in the present paper can also be used to obtain generalizations (in a somewhat different direction) pertaining to not necessarily positive and symmetric kernels; see [25] and [26, §7].

<sup>19</sup> See [18] or [31].

$x_0$  of  $E$  having the property described above. This number (which can be shown to be the same for all such neighborhoods  $U$ ) is called the (Leray-Schauder) index of  $g$  at  $x_0$ . It plays the role of the "multiplicity" of  $x_0$  as root of the equation (4.1). In particular,  $j \neq 0$  is a sufficient condition that  $x_0$  is a solution of (4.1), that is, a critical point. In the latter case,  $j$  is also called the index of the singularity  $x_0$  of the "vector field"  $g(x)$ . Under certain conditions the following relation<sup>20</sup> holds between the index  $j$  and the type numbers  $m^r$ :

$$(8.2) \quad j = \sum_r (-1)^r m^r.$$

Now let  $V = V_R$  be the solid sphere  $\|x\| \leq R$  of the Hilbert space  $E$ , and  $S = S_R$  its boundary. It follows easily from the complete continuity of  $G(x)$  that there are at most a finite number of critical points of the scalar  $i(x)$  in  $V_R$ . We denote these by  $a_\sigma$  ( $\sigma = 1, 2, \dots, s$ ), the index of  $g$  at  $a_\sigma$  by  $j_\sigma$ , and the  $r$ th type number of  $i(x)$  at  $a_\sigma$  by  $m_\sigma^r$ . We suppose that none of the critical points  $a_\sigma$  lies on  $S_R$ . Then under the mapping  $g$  the zero point  $0$  of  $E$  is not on the image of the boundary  $S_R$  of  $V_R$ , and therefore the mapping degree  $\chi = d(g, V_R, 0)$  is defined. This number is known to depend only on the behavior of  $g$  on  $S_R$ , and it is therefore also called the order of the image (under  $g$ ) of  $S_R$  with respect to  $0$ , or also the characteristic of the gradient field  $g(x)$  on  $S_R$ .<sup>21</sup> If we now apply (8.2) to each  $a_\sigma$  and sum in  $\sigma$ , we obtain immediately the relation

$$(8.3) \quad \chi = \sum_{\sigma=1}^s \sum_r (-1)^r m_\sigma^r$$

by applying the well known fact [21, Satz 5] that  $\chi = \sum_{\sigma=1}^s j_\sigma$ .

To show that this formula sometimes yields information about the number of critical points, we consider the case that the differential  $L(x, h)$  of  $G(x)$  exists, is completely continuous in  $h$ , and that moreover the second differential  $d^2(x, h, k)$  of  $i(x)$  is uniformly positive definite, that is, that there exists a positive constant  $\mu$  such that

$$(8.4) \quad d^2(x, h, k) \geq \mu \|k\|^2.$$

<sup>20</sup> For the proof in the finite-dimensional case, see [27]. While this seems to be the only published proof, the formula in this case was known independently to R. Bott and to M. Morse. For the proof in the Hilbert space case, see [28].

We remark (as R. Bott also did) that the well known theorem of Bendixson, according to which the index of an isolated singularity of a plane gradient field is never greater than 1, is a nearly immediate consequence of (8.2).

<sup>21</sup> For the concepts of order and characteristic, see [2, chap. 12] in the finite-dimensional case, and [21] for the Hilbert space case.

We claim that in this case  $V_R$  (for large enough  $R$ ) contains exactly one critical point, and that this critical point is a minimum [26, Theorem 6.2].

Indeed, that every eventual critical point  $a_\sigma$  of  $V_R$  is a minimum, and that  $m_\sigma^0 = 1$ ,  $m_\sigma^r = 0$  for  $r \geq 1$  follows immediately from (8.4). The relation (8.3) therefore specializes to

$$(8.5) \quad \chi = \sum_{\sigma=1}^s m_\sigma^0,$$

and since  $m_\sigma^0 \geq 1$  it will be sufficient to prove that

$$(8.6) \quad \chi = 1,$$

for then (8.5) implies  $s=1$ . To prove (8.6) we remark that since  $d^2(x, h, k)$  is the differential of the first differential  $d(x, h)$  of  $i(x)$ , we obtain from the mean value theorem [11, Theorem 5].

$$d(x, k) - d(0, k) = \int_0^1 d^2(tx, x, k) dt,$$

and therefore from (8.4), with  $k=x$ ,

$$(8.7) \quad \begin{aligned} d(x, x) &\geq \mu \|x\|^2 - |d(0, x)| \\ &= \|x\| \{ \mu \|x\| - |d(0, x)| / \|x\| \}. \end{aligned}$$

Recalling that, by the definition of a gradient in a Hilbert space  $E$  (last paragraph of §2),  $d(x, x) = (g(x), x)$ , we see immediately from (8.7) that to each large enough  $R$  there exists a positive  $\mu_1 = \mu_1(R)$  such that  $(g(x), x) \geq R^2 \mu_1(R) > 0$  for  $x$  on  $S_R$ . But this inequality means that the vector field  $g(x)$  on  $S_R$  is "exteriorly directed." Therefore its characteristic  $\chi$  on  $S$  is  $+1$  [21, part b of Satz 7].

As an application, we remark that the Hammerstein scalar (§3) is uniformly positive definite if, in addition to  $B$  and the assumptions made in §7, it satisfies the following condition:  $f$  has a first continuous derivative with respect to  $u$  satisfying the inequality  $(\partial f / \partial u) \geq -c$  where the constant  $c$  is subject to the condition that  $c^{-1}$  is greater than the greatest eigenvalue of the kernel  $K(s, t)$ . This implication follows easily from (3.13).<sup>22</sup> Consequently, the Hammerstein equation (3.12) has one and only one solution in this case.<sup>23</sup> Moreover, it follows from (3.12) that  $x(s) = -f(s, y(s))$  satisfies (3.11). Consequently, this  $x(s)$  minimizes the Hammerstein scalar.

<sup>22</sup> For details, see [26, §7].

<sup>23</sup> Concerning more general uniqueness statements, see the papers quoted in footnote 18.

## BIBLIOGRAPHY

1. L. Alaoglu, *Weak topologies of normed linear spaces*, Ann. of Math. vol. 41 (1940) pp. 252–267.
2. P. Alexandroff and H. Hopf, *Topologie*, Berlin, 1935.
3. N. Aronszajn, *The Rayleigh-Ritz and the Weinstein methods for approximation of eigen-values*, II: *Differential operators*, Technical Report 2, Oklahoma A. and M. College, Dept. of Math., Stillwater, Oklahoma, 1949.
4. S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932.
5. E. S. Citlanadze, *Certains problèmes de l'extrême relatif et de la théorie des valeurs caractéristiques*, C. R. (Doklady) Acad. Sci. URSS. vol. 56 (1947) no. 1.
6. ———, *Some problems of non linear operators and the calculus of variations in spaces of Banach type*, Uspehi Matematicheskikh Nauk N.S. vol. 5 (1950) pp. 141–142 (Russian; reviewed in Mathematical Reviews vol. 12 (1951) p. 110).
7. R. Courant and D. Hilbert, *Methoden der Mathematischen Physik*, vol. 1, 2d ed., Berlin, 1931; vol. 2, Berlin, 1937.
8. K. Friedrichs, *Spektraltheorie halbbeschränkter Operatoren und Anwendung auf die Spektralzerlegung von Differentialoperatoren*, Math. Ann. vol. 109 (1934) pp. 465–487.
9. M. Golomb, *Zur Theorie der nichtlinearen Integralgleichungen, Integralgleichungssysteme und allgemeinen Funktionalgleichungen*, Math. Zeit. vol. 39 (1934) pp. 45–75.
10. ———, *Ueber Systeme von nichtlinearen Integralgleichungen*, Publications Mathématiques de l'Université de Belgrade vol. 5 (1936) pp. 52–83.
11. L. M. Graves, *Riemann integration and Taylor's theorem in general analysis*, Trans. Amer. Math. Soc. vol. 29 (1927) pp. 163–177.
12. A. Hammerstein, *Nichtlineare Integralgleichungen und direkte Methoden der Variationsrechnung*, Berliner Math. Ges. Sitzungsber. vol. 26 (1927) pp. 66–77.
13. ———, *Nichtlineare Integralgleichungen nebst Anwendungen*, Acta Math. vol. 54 (1930) pp. 118–176.
14. E. Hille, *Functional analysis and semi-groups*, Amer. Math. Soc. Colloquium Publications, vol. 31, 1948.
15. M. Kerner, *Die Differentiale in der allgemeinen Analysis*, Ann. of Math. (2) vol. 34 (1933) pp. 546–572.
16. J. Leray and M. Schauder, *Topologie et équations fonctionnelles*, Ann. École Norm. (3) vol. 51 (1934) pp. 45–78.
17. L. Lusternik, *Sur une classe d'opérateurs non-linéaires dans l'espace de Hilbert*, Bull. Acad. Sci. USSR. Sér. Math. (1939) pp. 257–264.
18. M. Morse, *The calculus of variations in the large*, Amer. Math. Soc. Colloquium Publications, vol. 18, 1934.
19. B. J. Pettis, *A proof that every uniformly convex space is reflexive*, Duke Math. J. vol. 5 (1939) pp. 249–253.
20. F. Riesz, *Les systèmes d'équations linéaires à une infinité d'inconnues*, Paris, 1913.
21. E. H. Rothe, *Zur Theorie der topologischen Ordnung und der Vektorfelder in Banachschen Räumen*, Compositio Math. vol. 5 (1937) pp. 177–197.
22. ———, *Gradient mappings in Hilbert space*, Ann. of Math. vol. 47 (1946) pp. 580–592.
23. ———, *Completely continuous scalars and variational methods*, Ann. of Math. vol. 49 (1948) pp. 265–278.

24. ———, *Gradient mappings and extrema in Banach spaces*, Duke Math. J. vol. 15 (1948) pp. 421–431.
25. ———, *Weak topology and non-linear integral equations*, Trans. Amer. Math. Soc. vol. 66 (1949) pp. 75–79.
26. ———, *Critical points and gradient fields of scalars in Hilbert space*, Acta Math. vol. 85 (1951) pp. 73–98.
27. ———, *A relation between the type numbers of a critical point and the index of the corresponding field of gradient vectors*, Mathematische Nachrichten vol. 4 (1950/51) pp. 12–27.
28. ———, *Leray-Schauder index and Morse type numbers in Hilbert space*, Ann. of Math. vol. 55 (1952) pp. 433–467.
29. J. Schauder, *Ueber lineare elliptische Differentialgleichungen zweiter Ordnung*, Math. Zeit. vol. 38 (1934) pp. 257–282.
30. W. Schmeidler, *Integralgleichungen mit Anwendungen in Physik und Technik I*, Akademische Verlagsgesellschaft, Leipzig, 1950.
31. H. Seiffert and W. Threlfall, *Variationsrechnung im Grossen (Theorie von Marston Morse)*, Leipzig and Berlin, Teubner, 1938.
32. V. T. Soboleff, *Sur des éléments caractéristiques de certains opérateurs non linéaires*, C. R. (Doklady) Acad. Sci. URSS vol. 31 (1941) pp. 735–737.

UNIVERSITY OF MICHIGAN