fixed at the point $(0, 0)$ would, concerning relations (7), be a null vector.

Let us now make some simple (affine) transformation of the coordinates

(10) $$x = u, \quad y = u+v$$

where u, v are new variables. This transformation is allowable, since its jacobian becomes

(11) $$J = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1 \neq 0.$$

In the new coordinates the function σ will be expressed by a formula

$$\sigma = \frac{u(u+v)^2}{2u^2+2uv+v^2} = \varphi(u, v) \quad \text{for} \quad u^2+v^2 > 0.$$

Obviously

$$\varphi(h, 0) = h^3/2h^2 = h/2,$$

and so (since $\varphi(0, 0) = 0$ because for $u = v = 0$ we have $x = y = 0$)

$$\frac{\varphi(h, 0)-\varphi(0, 0)}{h} = \frac{1}{2}$$

or else

(12) $$\left(\frac{\partial \varphi}{\partial u}\right)_{u=v=0} = \frac{1}{2},$$

thus, in the system of coordinates (u, v), the vector fixed at the point $(0, 0)$ would have its first component different from zero, and thus it could not be a null vector. In this way the property of the vector to be a null vector would not be an invariant of the regular transformations of coordinates, and we have come to a contradiction in assuming $(\partial\sigma/\partial x, \partial\sigma/\partial y)$ to represent a vector field all over the plane.

References

[1] S. Gołąb, *Über den Begriff der Pseudogruppe von Transformationen*, Math. Ann. 116 (1939), p. 768-780.

[2] — *Sur une condition nécessaire et suffisante pour l'existence d'une différentielle totale*, Ann. Soc. Pol. Math. 16 (1937), p. 31-40.

Reçu par la Rédaction le 23. 2. 1956

ANNALES
POLONICI MATHEMATICI
VIII (1960)

On the notion of gradient

II. A certain extremal property of direction of the gradient vector

by S. Gołąb (Kraków) and M. Kucharzewski (Katowice)

If for a given scalar field σ (having at a point p under consideration a total differential or being at that point of the class C^1, i. e. having at that point continuous partial derivatives of the first order (see [1])) we form a gradient v whose components are the partial derivatives

(1) $$v_i = \partial\sigma/\partial x_i \quad (i = 1, \ldots, n),$$

then if the gradient v at the point p differs from zero, the direction of vector v defines that of the maximum increase of the field σ. By this the following is implied: If we take at the point p a fixed arbitrary radius r and its current point m, and if by s we denote the distance m from p and by $\varphi(s)$ the value of the function σ at the point m, then the right-side derivative of the function $\varphi(s)$ for $s = 0$ will be the greatest (and at the same time positive) for the radius r passing through the vector v. It is a well known fact and its proof is simple.

The aim of this note is to give some other similar property of the vector v.

About the function σ we assume that

1. it is continuous in the neighbourhood of the point p,
2. has at the point p a total differential,
3. the vector v with components

$$v_i = (\partial\sigma/\partial x_i)_p \quad (i = 1, \ldots, n)$$

is different from zero, i. e.

$$\sum_{i=1}^{n} |v_i| > 0,$$

4. there exists an infinite sequence of hypersurfaces S_ν with the following properties:

(I) S_ν is a closed $(n-1)$-dimensional surface containing p internally;

(II) denoting by β_ν the maximum distance of the current point of surface S_ν from p we have

(2) $$\lim_{\nu\to\infty}\beta_\nu = 0;$$

(III) defining by α_ν the minimum distance of the current point of surface S_ν from p we have

(3) $$\lim_{\nu\to\infty}\beta_\nu/\alpha_\nu = 1.$$

The function σ is by assumption continuous in the neighbourhood of p. S_ν is a compact set. It is contained in the neighbourhood of p for sufficiently great ν, and consequently σ attains in S_ν its greatest value, which will be denoted by μ_ν. Let y_ν be any of the points on the surface S_ν such that

(4) $$\sigma(y_\nu) = \mu_\nu.$$

Let us denote further by r_ν the radius issuing from p and passing through y_ν.

We assert that *under the above assumptions the sequence r_ν is convergent, and that the limit radius r_0 passes through the vector v.*

The above assertion says that if we close the point p by a hypersurface "approximate" to a hypersphere, and take any of the points of this hypersurface at which the function σ attains its maximum over the said hypersurface, then the direction of the vector from p to that point will be "near" to the direction of gradient v.

Proof. Without loss of generality we may assume that p is the origin of the system of coordinates and that

(5) $$\sigma(p) = 0.$$

Introducing short notations

(6) $$x \overset{\text{df}}{=} (x_1, \ldots, x_n), \qquad |x| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

and making use of the notion of scalar product

(7) $$x\cdot y = \sum_{i=1}^{n} x_i\cdot y_i$$

we may express the fact of possessing by the function σ the Stolz-Fréchet differential at the point p by means of the relation

(8) $$\lim_{|x|\to 0}\frac{\sigma(x) - x\cdot v}{|x|} = 0.$$

In order to prove our assertion it is sufficient to show that the angle φ_ν between the radius r_ν and the vector v tends to zero or, which means the same, that $\cos\varphi_\nu \to 1$, i. e.

(9) $$\lim_{\nu\to\infty}\frac{y_\nu\cdot v}{|y_\nu|\,|v|} = 1.$$

Let η be an arbitrary positive number. Let us choose for the number η a number ν_1 such that the inequality $\nu \geqslant \nu_1$ should imply the inequality

(10) $$\beta_\nu/\alpha_\nu < 1+\eta$$

as well as

(11) $$\alpha_\nu/\beta_\nu > 1-\eta/2.$$

This can be done in virtue of (3). Since by hypothesis we have $|v| > 0$, we can find for the number η a positive number ε such that

$$\varepsilon < \frac{\eta\,|v|}{4(1+\eta)}.$$

Subsequently, we shall choose for the number ε another number δ such that the inequality $|x| < \delta$ should imply the inequality

$$\left|\frac{\sigma(x) - x\cdot v}{|x|}\right| < \varepsilon,$$

which with respect to (8) is attainable.

Let us further choose a number ν_2 so that the inequality $\nu \geqslant \nu_2$ involves the inequality $\beta_\nu < \delta$, which can be done in virtue of (2).

Let $\nu_0 = \max(\nu_1, \nu_2)$ and let us assume $y \in S_\nu$, $\nu \geqslant \nu_0$. Then we have

$$|y| \leqslant \beta_\nu < \delta$$

and

$$-\varepsilon < \frac{\sigma(y) - y\cdot v}{|y|} < \varepsilon.$$

Hence

(12) $$y\cdot v - \varepsilon|y| < \sigma(y) < y\cdot v + \varepsilon|y| \quad \text{for} \quad y \in S_\nu \quad (\nu \geqslant \nu_0).$$

Let us denote by z_ν the point of intersection of S_ν with the vector v, and let us assume

$$\varrho_\nu \overset{\text{df}}{=} \max\{|y_\nu|, |z_\nu|\}.$$

Let us divide the hypersurface S_ν into two disjunctive sets A_ν and B_ν assigning to A_ν all those points y for which we have the inequality

$$y\cdot v < a_\nu|v| - 2\varepsilon\varrho_\nu,$$

and to B_ν the remaining points of the hypersurface S_ν. We assert that

$$y_\nu \in B_\nu.$$

For the indirect proof let us assume for a moment that $y_\nu \in A_\nu$. In this case

$$y_\nu\cdot v < a_\nu|v| - 2\varepsilon\varrho_\nu.$$

Hence by virtue of defining the number μ_ν and basing ourselves on inequality (12) we might have

$$\mu_\nu = \sigma(y_\nu) < y_\nu\cdot v + \varepsilon|y_\nu| < a_\nu|v| - 2\varepsilon\varrho_\nu + \varepsilon|y_\nu|.$$

However, $z_\nu \in S_\nu$, and thus $|z_\nu| \geqslant a_\nu$, whence we have

$$\mu_\nu < |z_\nu|\,|v| - 2\varepsilon\varrho_\nu + \varepsilon|y_\nu|.$$

It follows from the definition of the point z_ν that

$$z_\nu\cdot v = |z_\nu|\cdot|v|,$$

and consequently the last inequality may be re-written in the form

$$\mu_\nu < z_\nu\cdot v - 2\varepsilon\varrho_\nu| + \varepsilon|y_\nu|.$$

If we have

$$(13)\qquad |z_\nu| \geqslant |y_\nu|,$$

then $\varrho_\nu = |z_\nu|$, and consequently

$$\mu_\nu < z_\nu\cdot v - 2\varepsilon|z_\nu| + \varepsilon|y_\nu|.$$

On the other side, however, we have by (9):

$$z_\nu\cdot v - |z_\nu|\varepsilon < \sigma(z_\nu)$$

and since

$$\sigma(z_\nu) \leqslant \sigma(y_\nu) = \mu_\nu,$$

we have

$$z\cdot v - |z_\nu|\varepsilon < \mu_\nu < z_\nu\cdot v - 2\varepsilon|z_\nu| + \varepsilon|y_\nu|,$$

which leads to the conclusion

$$\varepsilon|z_\nu| < \varepsilon|y_\nu|,$$

i. e.

$$|z_\nu| < |y_\nu|,$$

contradictory to inequality (13).

If instead we have

$$|z_\nu| < |y_\nu|,$$

then $\varrho_\nu = |y_\nu|$ and

$$\mu_\nu = \sigma(y_\nu) < z_\nu\cdot v - 2\varepsilon|y_\nu| + \varepsilon|y_\nu| = z_\nu\cdot v - \varepsilon|y_\nu|$$
$$< z_\nu\cdot v - \varepsilon|z_\nu| < \sigma(z_\nu) \leqslant \mu_\nu,$$

and again we have a contradiction. The assumption that $y_\nu \in A_\nu$ leads to a contradiction, and therefore we have

$$y_\nu \in B_\nu$$

or else

$$y_\nu\cdot v \geqslant a_\nu|v| - 2\varepsilon\varrho_\nu.$$

Dividing the above inequality on both sides by $|y_\nu|\cdot|v|$ we obtain

$$\frac{y_\nu\cdot v}{|y_\nu|\cdot|v|} \geqslant \frac{a_\nu}{|y_\nu|} - \frac{2\varepsilon\varrho_\nu}{|y_\nu|\cdot|v|} \geqslant \frac{a_\nu}{\beta_\nu} - \frac{2\varepsilon\varrho_\nu}{|y_\nu|\cdot|v|}.$$

We have, however, $\varrho_\nu \leqslant \beta_\nu$, $|y_\nu| \geqslant a_\nu$, and this implies

$$\varrho/|y_\nu| \leqslant \beta_\nu/a_\nu$$

and furthermore

$$\cos\varphi_\nu = \frac{y_\nu\cdot v}{|y_\nu|\cdot|v|} > \frac{a_\nu}{\beta_\nu} - \frac{2\varepsilon}{|v|}\cdot\frac{\beta_\nu}{a_\nu}.$$

Now making use of inequalities (10) and (11) we may write

$$\frac{a_\nu}{\beta_\nu} - \frac{2\varepsilon}{|v|}\cdot\frac{\beta_\nu}{a_\nu} > 1 - \frac{\eta}{2} - \frac{2\varepsilon}{|v|}(1+\eta) > 1 - \frac{\eta}{2} - \frac{2(1+\eta)\eta|v|}{|v|4(1+\eta)}$$
$$= 1 - \frac{\eta}{2} - \frac{\eta}{2} = 1-\eta \quad \text{for} \quad \nu \geqslant \nu_0.$$

The last inequality proves that the sequence $\cos\varphi_\nu \to 1$, which was to be shown. Thus the theorem holds true.

The sequence of surfaces S_ν which, in the above theorem, shall satisfy conditions (I), (II), (III), cannot be replaced by a sequence of surfaces satisfying conditions (I) and (II) and instead of (III) a weaker one viz.

$$\text{(III}'\text{)} \qquad 0 < l < \beta_\nu/\alpha_\nu < L.$$

It is shown by the following counter-example.

Let us assume

$$\sigma(x, y) = x + y \qquad (n = 2).$$

The function σ thus defined satisfies all the assumptions of our theorem. Let us take the sequence of ellipses with the equations

$$\tfrac{1}{4}\nu^2 x^2 + \nu^2 y^2 = 1$$

as that of S_ν. Then the point y_ν, at which — as shown by a simple calculation — the function σ attains its maximum, has the coordinates $(4/\nu\sqrt{5}, 1/\nu\sqrt{5})$, and so the radius r_ν tends to the limit position (moreover, the radius r_ν in this example does not depend on ν)

$$x = \tau, \qquad y = \tfrac{1}{4}\tau, \qquad \tau \geqslant 0,$$

and consequently it does not possess the direction of the gradient which has the components $(1, 1)$.

The above theorem may serve as a generalization of the definition of direction for the gradient in the case where the gradient in the classical sense does not exist.

Indeed, let us encircle the point p by an arbitrary sequence of hyperspheres S_ν with the centre p and the radius tending to zero. In every hypersphere we arbitrarily choose a point y_ν where the function σ attains its maximum (or its upper limit). If the sequence of the radii r_ν has a definite limit position r_0 (independently of the choice of sequence S_ν and of the points y_ν) then *the direction r_0 is called the gradient direction of function σ at the point p.*

An easy example of the function σ defined as follows

$$\sigma(x, y) = \begin{cases} 0 & \text{for} \quad y \leqslant 0, \\ y & \text{for} \quad y > 0 \end{cases}$$

shows that at the points $p(x, 0)$ the function σ *does not possess a gradient but it has a definite gradient direction.*

Obviously, the existence of the gradient direction does not imply the value of the gradient (i. e. that which in the classical sense is the ength of the gradient vector).

If there exists at a given point p the gradient direction r_0, then the value of the generalized gradient (not necessarily absolute) may be defined as the limit

$$g = \lim_{m \to p} \frac{\sigma(m) - \sigma(p)}{\overline{mp}},$$

where m is a current point of the radius r_0 (if the classical gradient happens to exist, the value g is identical with $|v|$). It being thus, the sign of the value of the generalized gradient indicates whether the function σ increases or decreases in the direction of the gradient vector under consideration.

As shown by the following example, the value g need not exist. The example has been constructed for the simplest case $n = 2$.

Indeed, let us put

$$\sigma(x, y) = \varphi(y),$$

where the function $\varphi(y)$ is defined in the neighbourhood of zero and possesses the following properties: it is an increasing function, it has no right-side derivative for $y = 0$.

Let us consider the point $p(0, 0)$ and examine the existence of gradient direction. Let us describe a circle round the point p with the radius $\varepsilon > 0$. Let θ be the amplitude of a current point along the circumference. Then

$$\sigma(\varepsilon\cos\theta, \varepsilon\sin\theta) = \varphi(\varepsilon\sin\theta).$$

Since the function φ is by hypothesis an increasing one, we have

$$\varphi(\varepsilon\sin\theta) \leqslant \varphi(\varepsilon) \qquad \text{for} \qquad 0 \leqslant \theta \leqslant 2\pi.$$

Thus the function σ attains its maximum for $\theta = \pi/2$ and this fact is independent of the magnitude ε of the radius of the circle. Thus the existence of the gradient direction has been fully ensured, whereas the value g does not exist, because

$$\frac{\sigma(m) - \sigma(p)}{\overline{mp}} = \frac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon}$$

does not possess a limit for $\varepsilon \to 0 + 0$.

The example of function

$$\sigma = -\sqrt{x^2 + y^2 + \tfrac{1}{2}x^3}$$

determined in the neighbourhood of the point $p(0, 0)$ shows that the value g of the generalized gradient may be less than zero. Let us take for the above function the point $p(0,0)$. Let us describe a circle round this point with a radius ε. Then

$$\sigma(\varepsilon\cos\theta, \varepsilon\sin\theta) = -\varepsilon\sqrt{1+\tfrac{1}{2}\varepsilon\cos^3\theta}$$

attains the maximum value for $\theta = \pi$, viz.

$$\sigma(-\varepsilon, 0) = -\varepsilon\sqrt{1-\tfrac{1}{2}\varepsilon}.$$

Hence the gradient direction exists, whereas

$$g = \lim_{m\to p}\frac{\sigma(m)-\sigma(p)}{\overline{mp}} = \lim_{\varepsilon\to 0}\frac{-\varepsilon\sqrt{1-\frac{1}{2}\varepsilon}-0}{\varepsilon} = -1 < 0.$$

Remark 1. The generalized gradient indicates the direction of the maximum increase (or of the minimum decrease) of the function. Similarly the direction of the minimum increase (or that of the maximum decrease) could be introduced. If a gradient exists in the classical sense, then these two directions are in opposition. In general, it is not necessarily thus, and consequently there is some reason to speak about *two gradient directions*.

Remark 2. The above theorem holds true, as has been remarked by T. Ważewski, if the assumption 1 "the function is continuous in the neighbourhood of the point p" is replaced by a weaker one: "the function is defined in the neighbourhood of the point p". Then, evidently, the function σ need not attain its maximum in S_ν, but it has an upper limit in it, since the said function is bounded in the neighbourhood of the point p, owing to its possessing a differential at the point p. Then again, μ_ν would denote the upper limit of σ in S_ν, whilst by y_ν we should mean any of the points of S_ν with the following property:

There exists a sequence of points

$$q_{\lambda_\nu} \in S_\nu \quad (\lambda = 1, 2, \ldots)$$

convergent to y_ν, i. e. $\lim\limits_{\lambda\to\infty} q_{\lambda_\nu} = y_\nu$ such as

$$\lim_{\lambda\to\infty}\sigma(q_{\lambda_\nu}) = \mu_\nu.$$

The further statement of the theorem remains unchanged. The above change of the theorem would involve some, almost formal, changes in the argument.

Reference

[1] S. Gołąb, *On the notion of gradient. I. Essentiality of regularity suppositions*, this volume, p. 1-4.

Reçu par la Rédaction le 23. 2. 1956

ANNALES
POLONICI MATHEMATICI
VIII (1960)

On the notion of gradient

III. Gradient as a limit value of a surface integral

by S. Gołąb and A. Pliś (Kraków)

§ 1. Let a scalar field σ be given. If the function σ has at a definite point p a total differential, then at that point a gradient can be formed

$$v = \operatorname{grad}\sigma \tag{1}$$

as a vector with components (see [1])

$$v_i = \frac{\partial\sigma}{\partial x_i} \quad (i = 1, 2, \ldots, n). \tag{2}$$

We know the integral theorem which, under certain assumptions both about the field σ and about the closed hypersurface S bounding a finite and regular region D of space, expresses the integral of gradient over the region D by a surface integral. This theorem, containing in its vector form the so-called Green theorem, is stated thus:

$$\int_D \operatorname{grad}\sigma = \int_S N\cdot\sigma, \tag{3}$$

where N denotes the unit normal vector to S with an outside orientation.

To the above theorem corresponds the "differential" form, viz.

$$\operatorname{grad}\sigma(p) = \lim_{S\to p}\frac{1}{V}\int_S N\cdot\sigma. \tag{4}$$

V denotes here the volume (n-dimensional measure) of the region D.

Now the said formula is not precisely stated(1). We are concerned on one hand with assumptions with respect to the field σ and on the other with those referring to S, and finally with the limiting convergence

(1) It will suffice to see what W. Rubinowicz has written on this subject in his book [2], p. 67-70, in order to realize that the corresponding theorem is not stated in a satisfactorily strict way.