A GENERALIZATION OF SCHWARZ'S LEMMA*

BY CONSTANTIN CARATHÉODORY

1. *Introduction*. We consider the family of functions $f(z)$, which are regular inside of the unit circle, which vanish at the origin, and whose absolute value $|f(z)|$ is less than one in that circle. Taking two points z_1 and z_2 in the interior of the unit circle we inquire about the maximum $M(z_1, z_2)$ of the expression

$$\left|\frac{f(z_2) - f(z_1)}{z_2 - z_1}\right| \tag{1}$$

if $f(z)$ describes the family of functions considered above.

This maximum can never be less than one, because the function $f(z) \equiv z$ itself is contained among the functions of our family. But in a great number of cases $M(z_1, z_2)$ is *exactly equal to one*. Thus if z_1 is taken equal to zero, the assertion that $M(0, z_2) = 1$ is only another way of formulating the lemma of Schwarz. Again, if we assume that the ratio z_2/z_1 is real and negative, we have

$$|f(z_2) - f(z_1)| \leqq |f(z_1)| + |f(z_2)|,$$
$$|z_2 - z_1| = |z_1| + |z_2|;$$

and, using the lemma of Schwarz, we find that $M(z_1, z_2) = 1$.

In the third place, we have $M(z_1, z_2) = 1$ if *both* points z_1 and z_2 lie on the circular disc $|z| \leqq 2^{1/2} - 1$. This is an easy consequence of the fact that for all points of this figure the expression $|f'(z)|$ is never greater than one.† We are going to analyze the questions which arise from these different examples by determining completely all the cases for which $M(z_1, z_2) = 1$.

2. *An Auxiliary Function*. We begin with the obvious remark that our result will not be altered if we neglect from the outset all the functions of the form $f(z) = e^{i\theta} z$ for which the ex-

---

* From an address delivered before the Society under the title *Bounded analytic functions*, on November 27, 1936, by invitation of the Program Committee.

† J. Dieudonné, *Recherches sur quelques problèmes relatifs aux polynomes et aux fonctions bornées d'une variable complexe*, Annales de l'École Normale, (3), vol. 48 (1931), pp. 247–358; in particular, p. 352.

pression (1) is identically equal to one. For the remaining functions of our family we can put

$$f(z) = zg(z), \qquad \text{with} \qquad |g(z)| < 1, \tag{2}$$

and we may introduce the notation

$$g(z_1) = a. \tag{3}$$

The function $\phi(z)$ which is defined by the equation

$$\frac{a - g(z)}{1 - \bar{a}g(z)} = \frac{z_1 - z}{1 - \bar{z}_1 z}\phi(z), \tag{4}$$

is regular inside the unit circle; and it is readily seen that, for all points of that circle,

$$|\phi(z)| \leqq 1. \tag{5}$$

Using these formulas, we find

$$f(z) = z\frac{a(1 - \bar{z}_1 z) - (z_1 - z)\phi(z)}{(1 - \bar{z}_1 z) - \bar{a}(z_1 - z)\phi(z)}, \tag{6}$$

$$f(z_1) = z_1 a, \tag{7}$$

and finally

$$\frac{f(z_2) - f(z_1)}{z_2 - z_1} = \frac{a(1 - \bar{z}_1 z_2) + (z_2 - a\bar{a}z_1)\phi(z_2)}{(1 - \bar{z}_1 z_2) - \bar{a}(z_1 - z_2)\phi(z_2)}. \tag{8}$$

Now the function (6) always belongs to our family provided $|a|$ be taken less than one and $\phi(z)$ be a regular analytic function inside the unit circle which satisfies the condition (5). Consequently the value of $M(z_1, z_2)$ can be obtained by calculating the maximum of the absolute value of the right-hand side of (8) under the conditions

$$|a| < 1, \qquad |\phi(z_2)| \leqq 1. \tag{9}$$

For a given value of a, the maximum value of this last expression is attained at some point of the unit circle $|\phi| = 1$, say at the point $-e^{i\lambda}$. But if we multiply both a and $\phi(z_2)$ by $e^{-i\lambda}$, the absolute value of the right-hand side of (8) remains unaltered. Hence we may also calculate $M(z_1, z_2)$ as the maximum of $|\omega(a)|$ for $|a| < 1$, if we define $\omega(a)$ by the relation

$$\omega(a) = \frac{a(1 - \bar{z}_1 z_2) - (z_2 - a\bar{a}z_1)}{(1 - \bar{z}_1 z_2) + \bar{a}(z_1 - z_2)}, \tag{10}$$

which we obtain from (8) by putting $\phi(z_2) = -1$.

3. *A Necessary and Sufficient Condition.* We remark that we may write

$$\omega(a) = \frac{a(1 - \bar{z}_1 z_2 + \bar{a}z_1) - z_2}{(1 - \bar{z}_1 z_2 + \bar{a}z_1) - \bar{a}z_2}, \tag{11}$$

and that, with the notation

$$u = \frac{z_2}{1 - \bar{z}_1 z_2 + \bar{a}z_1}, \tag{12}$$

ω takes the form

$$\omega = \frac{a - u}{1 - \bar{a}u}. \tag{13}$$

The absolute value of this last expression is always greater than one for $|u| > 1$, equal to one for $|u| = 1$, and less than one for $|u| < 1$. This well known fact may be shown also by the formula

$$\frac{1 - |\omega|^2}{1 - |u|^2} = \frac{1 - |a|^2}{|1 - \bar{a}u|^2} > 0. \tag{14}$$

We have therefore the following result: if for some value of a (with $|a| < 1$) the value (12) of u has a modulus greater than one, we shall have, for this same value of a,

$$|\omega| > 1, \tag{15}$$

and consequently $M(z_1, z_2) > 1$ will hold. But if, for all values of $|a| < 1$, we always have $|u| \leqq 1$, it follows that we shall always have also $|\omega| \leqq 1$, and $M(z_1, z_2) = 1$. A necessary and sufficient condition that we shall have $M(z_1, z_2) = 1$ is therefore given by the inequality

$$|z_2| \leqq |1 - \bar{z}_1 z_2 + \bar{a}z_1|, \quad \text{for} \quad |a| < 1. \tag{16}$$

This being the case, the circle around the origin with radius $|z_2|$ has no point in common with the interior of the circle of center

$1-\bar{z}_1 z_2$ and radius $|z_1|$. Conversely, the inequality (16) will hold if these circles have no common point interior to both. Consequently the condition (16) may also be written in the form

$$(17) \qquad |z_1| + |z_2| \leqq |1 - \bar{z}_1 z_2|,$$

that is, in a form in which the parameter a is no longer involved.

4. *Simplification of the Condition.* This last inequality can be replaced by another which, although equivalent to it, is much simpler in form. Squaring both sides of (17), we get the condition

$$(18) \quad |z_1|^2 + |z_2|^2 + 2|z_1||z_2| \leqq 1 - \bar{z}_1 z_2 - z_1 \bar{z}_2 + |z_1|^2 |z_2|^2,$$

which is exactly equivalent to it. From the inequality

$$(19) \qquad - \bar{z}_1 z_2 - z_1 \bar{z}_2 \leqq 2|z_1||z_2|,$$

which is always true for all pairs of points z_1, z_2 for which (18) (or (17)) holds, it then follows that we must have

$$(20) \qquad (1 - |z_1|^2)(1 - |z_2|^2) \geqq 0.$$

It is therefore impossible that one of the points be outside and the other inside the unit circle, and a condition that none of these points be outside the unit circle is given by the inequality

$$(21) \qquad |z_1||z_2| \leqq 1.$$

Adding now to the members of (18) those of the identity

$$(22) \qquad |z_1 + z_2|^2 = |z_1|^2 + |z_2|^2 + z_1 \bar{z}_2 + \bar{z}_1 z_2,$$

and then reducing, we obtain

$$(23) \qquad |z_1 + z_2|^2 \leqq (1 - |z_1 z_2|)^2.$$

If (21) holds, this is equivalent to

$$(24) \qquad |z_1 + z_2| + |z_1 z_2| \leqq 1.$$

This last relation expresses therefore the condition that (17) and (21) hold simultaneously. Whenever it is satisfied, we need not state that the points z_1 and z_2 do not lie outside the unit circle. This most elegant form of the inequality (17) was pointed out to me by Szegö; it shows at first sight the symmetry in z_1 and z_2 of the condition obtained.

5. *Proof of Sufficiency of a Related Condition.* Take now two points z_1 and z_2 inside of the unit circle, for which (24) or the equivalent condition (17) holds; suppose also that

$$|a| \leqq \rho < 1. \tag{25}$$

We are going to show that under these assumptions the expression (1) cannot exceed a number which is actually less than one. We conclude first by the reasoning of §2 that any number which is not smaller than the upper bound of $|\omega(a)|$ under the condition (25) is suitable for our purpose.

Using (25) and (17), we have

$$\begin{aligned} |1 - \bar{z}_1 z_2 + \bar{a} z_1| &\geqq |1 - \bar{z}_1 z_2| - \rho |z_1| \\ &\geqq (1 - \rho)|z_1| + |z_2|, \end{aligned} \tag{26}$$

and consequently, by (12),

$$|u| \leqq \frac{|z_2|}{(1-\rho)|z_1| + |z_2|} = 1 - \frac{(1-\rho)|z_1|}{(1-\rho)|z_1| + |z_2|} < 1. \tag{27}$$

Both numbers $|a|$ and $|u|$ being not greater than one, we have the well known inequality

$$\begin{aligned} |\omega(a)| = \left| \frac{a - u}{1 - \bar{a}u} \right| &\leqq \frac{|a| + |u|}{1 + |a||u|} \\ &= 1 - \frac{(1 - |a|)(1 - |u|)}{1 + |a||u|}. \end{aligned} \tag{28}$$

Replacing in this inequality $|a|$ and $|u|$ by their upper bounds from (25) and (27), we get finally

$$|\omega(a)| \leqq 1 - \frac{(1-\rho)^2 |z_1|}{(1-\rho)|z_1| + (1+\rho)|z_2|}, \tag{29}$$

and we can therefore write

$$\left| \frac{f(z_2) - f(z_1)}{z_2 - z_1} \right| \leqq 1 - \frac{(1-\rho)^2 |z_1|}{(1-\rho)|z_1| + (1+\rho)|z_2|}. \tag{30}$$

6. *Statement of the Theorem.* We infer from this last result that the only functions $f(z)$ of our family for which, under the assumption of (24), the expression (1) attains its maximum value *one* are precisely those which we discarded at the beginning of

§2. This completes our proof and we can state the following theorem.

THEOREM. *For every pair of points* z_1, z_2 *lying inside the unit circle and satisfying the condition*

$$|z_1 + z_2| + |z_1 z_2| \leqq 1, \tag{31}$$

and for every analytic function $f(z)$ *which is regular for* $|z| < 1$, *which vanishes at the origin, and which fulfills the conditions* $|f(z)| < 1$ *everywhere in the circle, we always have*

$$\left|\frac{f(z_2) - f(z_1)}{z_2 - z_1}\right| < 1, \tag{32}$$

except for the case where $f(z)$ *is a linear function of the form* $e^{i\theta}z$.

For every pair of points z_1, z_2 *inside of the unit circle for which*

$$|z_1 + z_2| + |z_1 z_2| > 1, \tag{33}$$

there exist, on the contrary, analytic functions satisfying all of the above conditions for which the left-hand side of (32) *has values greater than unity.*

Taking $z_1 \neq 0$, and using the relation (30), we can replace the inequality (32) by another one that is more accurate. We remark for this purpose that we can take, in (30),

$$\rho = |a| = \frac{|f(z_1)|}{|z_1|}, \tag{34}$$

and that therefore we may write

$$\begin{aligned} 1 &- \left|\frac{f(z_2) - f(z_1)}{z_2 - z_1}\right| \\ &\geqq \frac{(|z_1| - |f(z_1)|)^2}{|z_1|(|z_1| - |f(z_1)|) + |z_2|(|z_1| + |f(z_1)|)}. \end{aligned} \tag{35}$$

This last inequality involves the following one, in which $|z_2|$ does not appear on the right side:

$$1 - \left|\frac{f(z_2) - f(z_1)}{z_2 - z_1}\right| > \frac{(|z_1| - |f(z_1)|)^2}{|z_1|(1 + |z_1|) + |f(z_1)|(1 - |z_1|)}, \tag{36}$$

$$(37)\qquad 1 - \left|\frac{f(z_2) - f(z_1)}{z_2 - z_1}\right| > \frac{(\,|z_1| - |f(z_1)|\,)^2}{2\,|z_1|}.$$

Finally, if we assume that $|z_2| \leqq |z_1|$, we can replace this last inequality by the following stronger one, which we obtain directly from (35):

$$(38)\qquad 1 - \left|\frac{f(z_2) - f(z_1)}{z_2 - z_1}\right| \geqq \frac{1}{2}\left(1 - \left|\frac{f(z_1)}{z_1}\right|\right)^2.$$

7. *The Geometric Meaning*. In order to find the geometric meaning of the condition (31), we take $z_1 = h$, where h is real, positive, and less than one; and we put $z_2 = x + iy$. We can then write

$$(39)\qquad ((x + h)^2 + y^2)^{1/2} \leqq 1 - h(x^2 + y^2)^{1/2}.$$

Squaring both sides of this last inequality, we get, after some reductions,

$$(40)\qquad 2h(x^2 + y^2)^{1/2} \leqq (1 - h^2)(1 - x^2 - y^2) - 2hx,$$

and this leads to the relation

$$(41)\qquad [(1 - h^2)(1 - x^2 - y^2) - 2hx]^2 - 4h^2(x^2 + y^2) \geqq 0.$$

Since the expression on the left side of (41) is positive at the origin as well as on circles $x^2 + y^2 = R^2$ for large values of R, and is negative on the circle

$$(42)\qquad (1 - h^2)(x^2 + y^2 - 1) + 2hx = 0,$$

the curve which is represented by

$$(43)\qquad [(1 - h^2)(x^2 + y^2 - 1) + 2hx]^2 - 4h^2(x^2 + y^2) = 0$$

consists of two loops, one of which lies inside of the circle (42) and the other outside of the same circle. On the unit circle $x^2 + y^2 = 1$, the left side of (43) has the value

$$4h^2(x^2 - 1) \leqq 0,$$

which shows that the unit circle also divides both loops.

It follows finally that the inner loop of the curve (43) divides the unit circle into two regions. In the one of these regions, which is convex, the condition $|z + h| + |zh| < 1$ is fulfilled. The boundary of this region, which we shall call A, touches the unit circle at the point $x=-1$, $y=0$. In the second region B, which is horn-shaped, we have $|z+h|+|zh|>1$.

8. *The Curve whose Stereographic Projection is the Curve* (43). Since the curve (43) is bicircular, that is, since it has the circular points at infinity as double points, it is convenient to consider it as the stereographic projection of a spherical curve. The stereographic projection of the sphere

$$\xi^2+\eta^2+\zeta^2=\frac{1}{4} \tag{44}$$

on the plane with the coordinates x and y is defined by the formulas

$$x=\frac{\xi}{\frac{1}{2}-\zeta}, \qquad y=\frac{\eta}{\frac{1}{2}-\zeta}, \qquad x^2+y^2=\frac{\frac{1}{2}+\zeta}{\frac{1}{2}-\zeta}. \tag{45}$$

With these new variables, the equation (43) takes the form

$$h^2\xi^2+2h(1-h^2)\xi\zeta+(1-h^2+h^4)\zeta^2=\frac{h^2}{4}. \tag{46}$$

The spherical curve with which we have to deal consists therefore of the intersection of the sphere (44) with the elliptic cylinder (46).

We introduce the new rectangular coordinates ξ' and ζ' by the formulas

$$\xi=\xi'\cos\phi+\zeta'\sin\phi, \qquad \zeta=-\xi'\sin\phi+\zeta'\cos\phi, \tag{47}$$

$$\cos\phi=\frac{1}{(1+h^2)^{1/2}}, \qquad \sin\phi=\frac{h}{(1+h^2)^{1/2}}. \tag{48}$$

The equation (46) then takes the form

$$h^2\xi'^2+\frac{1}{h^2}\zeta'^2=\frac{1}{4}. \tag{49}$$

The part of this curve which corresponds to the inner loop of the bicircular curve (43) is determined by the inequalities

$$-\frac{1}{2(1+h^2)^{1/2}}\leqq\xi'\leqq\frac{1}{2(1+h^2)^{1/2}}, \qquad \zeta'\leqq\frac{-h}{2(1+h^2)^{1/2}}. \tag{50}$$

9. *Osculating Circles at the Vertices of the Curve* (43). Using these formulas, we can obtain readily all of the properties of the curve (43) which are important for the problem we are considering. Thus we can show that, when h increases from zero to one, the region A shrinks continuously; for small values of h, this region fills nearly the whole of the unit circle; if h tends towards one, this region reduces to a narrow band which surrounds the radius extending from -1 to the center.

As an example of such computations we shall determine the osculating circles at the four vertices of the boundary of the region A, that is, at the four points in which the osculating circles do not cross the curve.

Two of these points correspond to the end-points of the arc (50) of the ellipse (49). If we determine the tangents to the ellipse at these points and the segments of these tangents lying inside the circle

$$\xi'^2 + \zeta'^2 = \frac{1}{4}, \qquad \eta = 0, \tag{51}$$

the stereographic projections of these segments coincide with diameters of the osculating circles for which we are looking. We find in this way that the diameters of these circles are given by the formulas

$$-1 \leqq x \leqq \frac{1 - h - h^2}{1 + h - h^2}, \qquad y = 0, \tag{52}$$

$$-\frac{1 - h^3}{1 + h^3} \leqq x \leqq \frac{1 - h}{1 + h}, \qquad y = 0. \tag{53}$$

We can also write the equations of the circles themselves; we find

$$(1 + h - h^2)(x^2 + y^2) + 2hx - (1 - h - h^2) = 0 \tag{54}$$

for the circle with the diameter (52), and

$$\begin{aligned} &(1 + h^3)(x^2 + y^2) + 2h(1 - h)(1 + h^3)x \\ &\qquad\qquad - (1 - h)^2(1 + h^2 + h^4) = 0 \end{aligned} \tag{55}$$

for the other circle.

The two other vertices of the bicircular curve correspond to the vertex $\xi'=0$, $\zeta'=-h/2$ of the ellipse (49), and are symmetric with respect to the x axis. The osculating circle passing through that point is the stereographic projection of a circle on the sphere. Putting

$$u = \zeta' + \frac{h}{2} \tag{56}$$

and determining two suitable constants p and q, we see that this circle lies on the plane

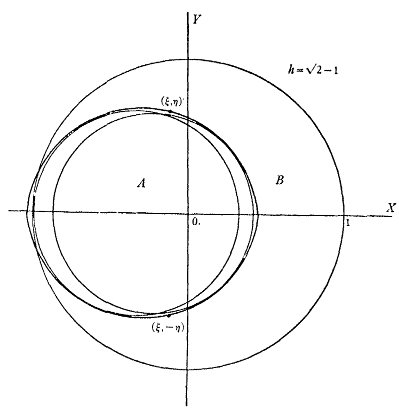

FIG. 1

$$2\eta = pu + q. \tag{57}$$

Using (56), we can write the equation of the ellipse (49) in the form

$$4h^4\xi'^2 - 4hu + 4u^2 = 0, \tag{58}$$

and we see that the intersection of the sphere (44) with the plane (57) has the form

$$4\xi'^2 + (pu + q)^2 + (2u - h)^2 = 1. \tag{59}$$

The curves (59) and (58) osculate at the point $u=0$ if we have

$$q = \pm (1 - h^2)^{1/2}, \qquad - h^4pq = 2h(1 - h^4). \tag{60}$$

This gives for the equation of the two circles of osculation in the original variables

$$
(61) \quad \begin{aligned} &(1 - h^2)^{1/2}(h(x^2 + y^2 + 1)) \\ &\qquad + (1 + h^2)^{1/2}(2hx + (x^2 + y^2 - 1)) \pm 2h^3y = 0. \end{aligned}
$$

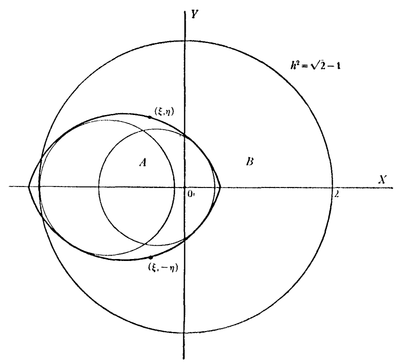

FIG. 2

It is very easy to sketch the bicircular curve which forms the boundary of the region A after the circles (54), (55), and (61) have been drawn; indeed, both circles (54) and (55) lie completely inside the region A. Both circles (61) surround A and a rather large arc on each of these circles may be considered as coinciding with the boundary of A. In Figs. 1 and 2, these circles have been drawn first for the case $h=2^{1/2}-1$, in which the point $z=h$ lies on the boundary of the region A, and secondly for $h^2=2^{1/2}-1$, in which case the circles (61) have radii of minimum value.

THE UNIVERSITY OF WISCONSIN