A Topological Generalization of
Schwarz's Lemma

by
George Piranian

A thesis presented to the Faculty of The Rice Institute in partial fulfillment of the requirements for the Master of Arts degree.
The classical lemma of Schwarz states that if \( f(z) = zg(z) \) where \( g(z) \) is holomorphic inside the unit circle, and if \( |f(z)| \leq 1 \) when \( |z| < 1 \), then \( |f(z)| \leq |z| \) when \( |z| < 1 \). The proof for this is based on the fact that if \( f(z) \) is holomorphic in a region, its absolute value has no relative maxima in the interior of the region.

Max Zorn* has stated and proved a more general, highly axiomatic version of Schwarz's lemma, applicable to certain families of transformations of a metrizable topological space into itself. But a geometrical interpretation of Zorn's results is impossible without severe additional restrictions on the space and the families of transformations.

The present paper proves a geometrical version of Schwarz's lemma by metrical-topological methods. The proof is applicable to the euclidean \( n \)-space in particular, and, more generally, to any convex topological space provided each point of the space lies on one of the surfaces of a system of concentric compact spheres.

Notation: F, G, H, I, P, R denote transformations; in particular, I is the identity transformation, P transforms the space into a point, and R is a rotation. x, y, p, q are points. The positive integers m, n, i, when used as subscripts, denote a sequence of transformations, points, or numbers; sometimes the subscript i is used to refer to a single term from a sequence. $F^1, x^n, n^i$ are order-preserving subsequences of $F^n, x^n, n^i$. $F^n$ denotes either the iterated transformation F or the sequence of such transformations.

$|x|$ is the distance from the origin p to the point x. Brackets about the serial number of a statement indicate a definition or a postulate, parentheses a theorem.
1. Preliminary definitions and theorems.

1.1. \( S \) is a convex metrized topological space with the property that the surface of every sphere in \( S \) is a compact set. To avoid triviality, we shall assume that \( S \) contains more than one point.

1.2. A sequence \( F_n \) of transformations of \( S \) into itself converges to \( F \) if \( \lim x_n = x \) implies \( \lim F_n(x_n) = F(x) \).

1.3. \( F_n \) is properly divergent if \( F_n(x) \) fails to converge for every point \( x \).

1.4. The product \( H = FG \) is defined by \( H(x) = FG(x) = F(G(x)) \).

1.5. If \( \lim F_n = F \) and \( \lim G_n = G \), then \( \lim F_n G_n = FG \).

Proof: Suppose \( \lim F_n = F, \lim G_n = G, \) and \( \lim x_n = x \). Then

\[ \lim G_n(x_n) = G(x), \text{ and therefore} \]

\[ \lim F_n(G_n(x_n)) = F(\lim G_n(x_n)) = FG(x). \]

1.6. If \( \lim F_n = F \) and \( \lim F_n^{-1} = G \), then \( G = F^{-1} \).

Proof: \( GF = (\lim F_n^{-1})(\lim F_n) = \lim F_n^{-1}F_n = I. \)

1.7. If \( F \) maps every point on the same point \( p \), \( F \) is a constant, and we call it \( P \).

1.8. \( F \) is nilpotent if \( \lim F^n \) is a constant.

1.9. If \( F \) is nilpotent, \( F \) has exactly one fixed point.

Proof: If \( \lim F^n(x) = p \), then \( F \lim F^n(p) = \lim F^{n+1}(p) = p \), so that \( F(p) = p \). If on the other hand \( \lim F^n(x) = p \) and \( F(q) = q \), then \( \lim F^n(q) = q \), and since \( \lim F^n(x) = p \) for all \( x \), \( p = q \), and the fixed point is unique.
2. Description of the normal family $\mathcal{N}$.

2.1. The family $\mathcal{N}$ is any set of transformations with the following properties:

2.1.1. The elements of $\mathcal{N}$ are single-valued (many-one) continuous transformations of $S$ into all or a part of $S$. 

2.1.2. $\mathcal{N}$ contains $I$, and it contains the product of any two of its elements.

2.1.3. If $F$, $G$, $H$ are in $\mathcal{N}$, and if $F$ is not a constant, then $GF = HF$ implies $G = H$.

2.1.4. Every sequence $F_n$ in $\mathcal{N}$ contains a subsequence $F'_n$ which is either convergent to some $F$ in $\mathcal{N}$, or else properly divergent.

2.1.4.1. It is of interest to compare the family $\mathcal{N}$ with its analogue in the theory of analytic functions. In function theory, a family $\mathcal{N}$ is said to be normal in $|x| < 1$ if every sequence $F_n$ in $\mathcal{N}$ contains a subsequence $F'_n$ which converges uniformly to some $F$ (not necessarily in $\mathcal{N}$) or diverges uniformly to $\infty$ in every closed region inside $|x| < 1$.

In (2.1.4.2) we shall show that if a sequence $F_n$ in our normal families $\mathcal{N}$ converges to $F$ in $\mathcal{N}$, it does so uniformly in every bounded, closed, compact set in $S$, so that the condition in function theory is satisfied.

If $F_n$ is properly divergent, either there exists a subsequence $F'_{n'}$ that converges to some $F$ in $\mathcal{N}$ (and therefore converges uniformly in any bounded, closed, compact set), or every subsequence of $F_n$ is properly divergent. In the latter case, $|F_n(x_{0})| \to \infty$ for every $x_0$; but it does not follow a priori that this divergence to $\infty$ is uniform in every bounded, closed, compact set; i.e., the conditions in function theory are not necessarily satisfied.

(2.1.4.2) Suppose that the convergence of $F_n$ in $\mathcal{N}$ to $F$ in $\mathcal{N}$ is not uniform in some bounded, closed, compact set $S_\epsilon$ in $S$. Then there...
exists an \( r_0 > 0 \) and a sequence of points \( x_n \) such that \( x_n \to x \) in \( S_p \) and \( \left| F_{x_n}(x_n) - F(x_{x_n}) \right| > r_0 \), and since \( \lim F(x_{x_n}) = F(x) \),
\[ \left| F_{x_n}(x_n) - F(x) \right| > \frac{r_0}{2}, \]
in contradiction with the fact that \( \lim F_{x_n} = F \).

2.2. The transformation \( F \) in \( N \) is a rotation \( R \) about \( p \) if
\[ |F(x)| = |x| \text{ for all } x. \]

2.2.1. \( L_x \) is the surface of the sphere with center \( p \) and radius \( |x| \).

The component of \( S - L_x \) which contains \( p \) is the interior of the sphere. If \( F \) is a topological transformation, that part of \( S - F(L_x) \) which contains \( F(p) \) is the interior of the surface \( F(L_x) \).

2.3. The family \( H \) shall also satisfy the following conditions:
\( p \) is a certain fixed point, called the origin, and if \( y \) lies in \( L_x \), \( N \) contains an \( R \) about \( p \) with the property \( R(x) = y \).

2.4. Henceforth, whenever a transformation is designated by \( F \) or by \( R \), with or without subscript or accent, it shall be assumed that the transformation is in the family \( N \) under discussion. It shall also be assumed that \( F(p) = p \) holds for these transformations. This does not mean that \( p \) is fixed under every transformation in \( N \), but merely that our conclusions apply to transformations in \( N \) for which \( F(p) = p \).

Since all functions constructed in proofs will be rotations or limits of sequences \( F_n \), where \( F_i(p) = p \) for all \( i \), the assumption \( F(p) = p \) will always be justified.

3. Topological properties of elements in \( N \).

3.1. If \( F_\alpha(x_\alpha) \) is compact (the index \( \alpha \) ranging over an arbitrary set of symbols), then \( F_\alpha(x) \) is compact for all \( x \).

Proofs: Every sequence from \( F_\alpha(x_\alpha) \) contains a subsequence \( F_{\alpha'}(x_{\alpha'}) \) converging to some point \( q_\alpha \), and therefore every sequence from \( F_\alpha(x) \) contains a convergent subsequence, by [2.1.4]. Therefore \( F_\alpha(x) \) is
compact for all $x$.

(3.1.1) Every sequence $F_n$ has a convergent subsequence $F_{n'}$.

(3.2) If $F_{n'}$ tends to the constant $F_1$ ($n_{i+1} > n_i$), $F$ is nilpotent and $\lim F' = p$.

Proof: If $\lim F_{n'} = P$, every sequence $F_{m'}$ ($m_{i+1} > m_i$) contains a subsequence converging to $P$, for from $m_i$ and $n_i$ we can select subsequences $m_i'$ and $n_i'$ such that $d_i' = m_i' - n_i'$ is increasing and $F_{d_i'}$ is a convergent sequence with $\lim F_{d_i'} = F'$. Since $F^{n_i'} = F_{n_i'} F_{d_i'}$ for all $i$, $\lim F^{n_i'} = \lim F_{n_i'} F_{d_i'} = \lim F^{n_i} F' = PP' = P$, and the subsequence from $F^{n_i'}$ has been found.

Now suppose that the theorem is false. Then there exists a convergent sequence $x_i$ such that $F^{n_i'}(x_i)$ does not converge to $p$; and therefore there exist sequences $m_i'$, $x_i'$ and a positive constant $r$ such that $|F^{n_i'}(x_i)| > r$ for all $i$, contrary to what we have just proved.

(3.3) If a convergent sequence $F_{n'}$ with the non-constant limit $F'$ exists, then $F$ has an inverse in $N$.

Proof: We select the subsequence $n_i'$ such that $d_i' = n_i' - n_i - 1$ is strictly increasing and $F_{d_i'}$ is a convergent sequence with $\lim F_{d_i'} = F'$. Then

$$\lim F_{n_i} = \lim F_{n_i+1} = \lim F_{n_i'} F_{d_i'} F = (\lim F_{n_i'})(\lim F_{d_i'}) F$$

i.e., $F^* = F^* F = F^* F' F$, and by [2.1.3], $F' F = I$, and $F' = F^{-1}$.

(3.3.1) If $F$ is not nilpotent, it is a topological transformation; i.e., it is bi-continuous.

(3.3.2) Every rotation has an inverse.

Proof: If $\lim x_{n_i} = x \neq p$, then $\lim |R^n(x_{n_i})| = |x| \neq 0$, and $R$ is not nilpotent.

(3.4) If $F$ is nilpotent, $F^{-1}$ is not in $N$.

Proof: Suppose that $F$ is nilpotent and has an inverse in $N$. Since $F^{-1}(p) = p$, a convergent sequence $F^{-n_i}$ exists. This implies

$I = \lim F^{-n_i} F_{n_i} = (\lim F^{-n_i}) P$, which is absurd.
4. Schwarz' Lemma

\[ 4.1 \]
\[ D(x)F = \max_{|y| = |x|} |F(y)| - |y|, \]
\[ d(x)F = \min_{|y| = |x|} |F(y)| - |y|. \]

(4.1.1) \( D(x)F \) and \( d(x)F \) are continuous functions of \( x \).

**Proof:** Suppose \( \lim_{x \to x_0} = x \) where \( |F(x)| = |x| = D(x)F \). Then
\[ \lim |F(x_0)| = |x| + D(x)F, \]
and therefore
\[ \lim D(x_0)F \geq D(x_0). \]

To show that \( \lim D(x_0)F \leq D(x_0)F \), we suppose the contrary to be the case. Then there exists a sequence \( y_\alpha \) with \( \lim y_\alpha = y \) and \( |y| = |x| \) such that
\[ D(y_\alpha)F = |F(y_\alpha)| = |y| > D(x)F + k > |F(y)| - |y| + k \]
(where \( k \) is some positive constant), so that \( F \) cannot be continuous.

From the two inequalities it follows that \( D(x)F \) is continuous.

The proof for \( d(x)F \) is similar.

(4.1.2) For all \( R_1 \) and \( R_2 \), \( D(x)F = D(x)R_1 FR_2 \) and \( d(x)F = d(x)R_1 FR_2 \).

**Proof:** Since \( |R_1(y)| = |y| \) and \( |R_2F(y)| = |F(y)| \) for all \( y \),
\[ \max_{|y| = |x|} |F(y)| = |y| = \max_{|y| = |x|} |R_1 FR_2(y)| - |y|, \]
\[ \min_{|y| = |x|} |F(y)| - |y| = \min_{|y| = |x|} |R_1 FR_2(y)| - |y|. \]

(4.2) The necessary and sufficient condition for \( F \) to be nilpotent is \( D(x)F < 0 \) for all \( x \ (x \neq p) \).

**Proof:** Suppose \( F \) is nilpotent; then \( d(x)F < 0 \) for some \( x \); for otherwise we should have \( |F(x)| \geq |x| \) for all \( x \), and therefore
\[ |F^n(x)| \geq |x| \]
for all \( n \) and all \( x \), contrary to hypothesis.

Suppose now that for some \( x \ (x \neq p) \) either \( d(x)F = 0 \), or \( D(x)F = 0 \), or \( D(x)F \) and \( d(x)F \) are of opposite sign. Then, for some \( y \) in \( L_x \), \( |F(y)| = |y| \), and there exists an \( R \) with the property
RF(y) = y, so that RF has two fixed points and cannot be nilpotent.

Then (RF)^{-1} is in N, and it follows that F has the inverse F^{-1} = (RF)^{-1} R,
since (RF)^{-1} FF = I. Therefore F is not nilpotent, and the necessity of
the condition is proved.

To prove that the condition is also sufficient, suppose that
D(x)F < 0 for all x (x \neq p). Let E(r, R) be the set of all points x
with the property r \leq |x| \leq R < \infty. Then, by continuity of D(x)F, we
have for every E(r, R) (0 < r, R < \infty |x| in S) a value k < 0 such
that D(x)F < k when x is in E(r, R). It follows that if -nk > R - r,
F^{-n}(E(r, R)) \subseteq E(0, r), so that F is nilpotent.

(4.8) If F has an inverse in N, d(x)F \geq 0 for all x.

Proof: Suppose that, contrary to the theorem, F has an inverse and
d(x_0)F < 0. We construct the sequence F\_n:

F\_n = FR_1 FR_2 \ldots FR_n,

where the set R \_n is any set satisfying the inductive definition
R \_n F(x\_i) = x\_i+1, x\_i+1 being any point such that

|F(x\_i+1)| - |x\_i+1| = d(x\_i+1)F\_n.

First we shall show that d(x\_i)F < 0 for all i. Suppose the
contrary; let i be the smallest integer such that d(x\_i)F \geq 0, and choose
x\_i so that |x\_i+1| > |x\_i| > |x\_i|. Since the sphere L_{x\_i} divides S
into two disjoint sets: E\_1 (|x| \leq |x\_i|) and E\_2 (|x| > |x\_i|), and
since R\_i, F has an inverse, so that topological properties of sets are
invariant under R\_i, F, we have x\_i \equiv R\_i, F(x\_i) \subseteq R\_i, F(E\_2). But
also, since R\_i, F(E\_1) \subseteq R\_i, F(E\_1) and R\_i, F(L_{x\_i}) has no points interior
to L_{x\_i} (d(x\_i)F \geq 0), we have x\_i \subseteq R\_i, F(E\_1), so that the hypothesis
d(x\_i)F \geq 0 is untenable.

Therefore the sequence |F\_n(x\_0)| is decreasing, and the
limit points of F\_n(x\_0) all lie on a sphere L_{x\_i}. But \lim d(x\_i)F = 0,
and by the continuity of \( d(x)F \), \( d(y)F = 0 \). Suppose \( F^n \) has the limit \( F^* \), and \( F^* (x_0) = y \). Then \( d(y)F^* = 0 \). For \( d(y)F = 0 \), and therefore \( d(y)F^2 = 0 \), since \( F \) is a topological map. By induction, \( d(y)F_i = 0 \) for all \( i \), and \( \lim d(y)F_i = d(y)F^* = 0 \).

Therefore \( F^* (x_0) \) lies on \( F^* (L_\omega) \) or interior to it, so that \( F^* \) cannot be a topological transformation. But this is impossible; for \( F \) has an inverse, and \( \{ d(y)F^* \} \) is a sequence converging to \( F^* \) in \( \mathbb{N} \). Therefore the supposition \( d(x_0)F < 0 \) is untenable if \( F \) has an inverse in \( \mathbb{N} \).

(4.4) Either \( F \) is nilpotent, or \( F \) is a rotation.

Proof: If \( F \) is not nilpotent, then \( d(x)F \geq 0 \). But if \( |F(x)| < |x| \), then \( d(y)F^{-1} < 0 \) where \( y = F(x) \), and it follows that \( F^{-1} \) has no inverse in \( \mathbb{N} \), which is absurd. Therefore \( d(x)F = D(x)F = 0 \) for all \( x \), and \( F \) is a rotation.

5. Two theorems concerning the rotations in \( \mathbb{N} \).

In this section, we assume that \( S \) is contained in the euclidean plane or its topological equivalent, and that there is a system of coordinates \( r, t \) on \( S \) \((0 \leq t < 1)\), with the properties

(i) if \( x \equiv (r, t) \), \(|x| = r^2 \)

(ii) \( t \) is a parameter, continuous (modulo 1) throughout \( S \), except at the origin.

If \( R \) maps the point \((r, t)\) on \((r, f(t))\), we shall write \( R(t) = f(t) \).

(5.1) If \( \mathbb{N} \) contains every rotation of the form \( R(t) = t + k \) \((k \) any real number, \( t + k \) reduced modulo 1), then every rotation in \( \mathbb{N} \) is of this form, unless it is a reflection.

Proof: Suppose \( \mathbb{N} \) contains all rotations \( R_k (t) = t + k \), and suppose \( R \) is a rotation, other than a reflection, with the property \( R(0) = k \), \( R(h_0) \neq h_0 + k \) for a certain \( r \). Then \( R_k (0) = 0 \), \( R_R (h_0) = h_1 \neq h_0 \).
If we write \((R_k R)^\infty(h_0) = h\) and assume \(h\) cannot lie between 0 and \(h\), \(h_\infty\) for \(R_k R\) transforms the set \(0, h_\infty, h_1\) into \(0, h_1, h_2, \ldots\) and the supposition that \(R_k R\) alters the relative position of \(0, h_\infty, h_1\) on the circle is contrary to the fact that \(R\) is not a reflection. By induction, \(h_\infty, h_1, h_2, \ldots\) is a strictly monotonic sequence, which must converge, since \(R_k R(0) = 0\). If \(\lim h_\infty = h\) and \((R_k R)^\infty\) converges to \(R^*\) in \(\Omega\), \(R^*(h) = h = R^*(h_0)\). Therefore \(R^*\) cannot be a topological map, and \(R\) is not in \(\eta\).

It follows that every rotation in \(\eta\) is of the form \(R(r, t) = (r, t + k(r))\). It remains to be shown that, for every \(R\), \(k(r)\) is constant.

Suppose \(R(r, t) = (r, t + k(r))\), where \(k(r)\) is not constant. Clearly, \(k(r)\) is continuous. Let \(r_\infty\) be an increasing sequence converging to \(r'\), with the property \(k(r_\infty) \neq k(r')\). In the interval \([r_\infty, r']\), \(k(r)\) takes on every value in \([k(r_\infty), k(r')]\). Let \(C_{r_\infty}\) be the curve \(t = 0, r_\infty \leq r \leq r'\). If \(n > \frac{1}{|k(r_\infty) - k(r')|}\) and \(t_0\) is any number in \([0, 1]\), the curve \((R_{r_\infty}^{(n)})(C_{r_\infty})\) has at least one point with the coordinate \(t_0\). Suppose \(\lim R_{r_\infty}^{(n)} = R^*\). Then \(\lim R_{r_\infty}^{(n)} C_{r_\infty} = L_{r'}\), where \(L_{r'}\) is the circle \(r = r'\). Therefore \(R^*\) is not continuous, and \(R\) cannot be in \(\eta\).

(5.2) The rotation \(R_k(t) = t + k\) is contained in exactly one family \(\eta\) of rotations, or in at least two distinct families \(\eta_1, \eta_2\), depending on whether \(k\) is irrational or rational.

**Proof:** If \(k\) is irrational, the set \(\lambda = k n_\infty (n = 1, 2, \ldots; k n \text{ reduced modulo 1})\) is everywhere dense in \([0, 1]\), and by [2.1.4] the family \(\eta\) containing \(k\) must contain all rotations of the form \(R_k(t) = t + h\) (\(h\) any real number), and it is thereby uniquely determined.

If \(k\) is rational (say \(k = \frac{m}{n}\)), \(R \Rightarrow(t) = t + k\) is contained in the family \(\eta_1\), composed of all the rotations \(R_k(t) = t + h\). We construct the family \(\eta_2\) so that \(\eta_2\) contains \(R_k\), and \(\eta_1, \eta_2\) have only \(n\) rotations in common.
If we write \( \varphi(t) = t + \frac{1}{2 \pi} \sin 2 \pi t \), \( \varphi(t) \) is a strictly increasing function, and therefore it has a strictly increasing inverse \( \varphi^{-1} \). Therefore \( R^\varphi \), defined by
\[
R^\varphi(t) = \varphi^{-1} R^\varphi \varphi(t)
\]
is a topological transformation. Since
\[
R^\varphi R^\varphi(t) = \varphi^{-1} R^\varphi \varphi \varphi^{-1} R^\varphi \varphi(t) = \varphi^{-1} R^\varphi \varphi(t) = R^\varphi \varphi(t) = R^\varphi \varphi(t) \overline{2 + g(t)}
\]
the set \( R^\varphi \) (\( h \) assuming all real values) constitutes a normal family \( N \). Moreover,
\[
R^\varphi(t) = \varphi^{-1}(t + \frac{1}{2 \pi} \sin 2 \pi t)
\]
so that the family \( N \) contains \( R^\varphi \); and since \( \varphi(t) \) has the primitive period \( \frac{1}{n} \), \( N \) contains no members of \( N \), except those generated by \( R^\varphi \).
Comparison of the present method with that of Zorn.

The work in this paper, up to [2.1.4] inclusively, follows that of Zorn, except that Zorn assumes of $S$ only that it is a metrizable topological space.

Zorn's simplest version of Schwarz' lemma requires no more of $N$ than that it satisfy the conditions in [2.1]. Zorn defines that a transformation $R$ in $N$ is a rotation provided $R(p) = p$ and $R$ has an inverse in $N$. He defines circumferences in terms of the normal family $N$ as follows: the set $L_x$ containing the point $x$ is a circumference provided

(i) $R(L_x) = L_x$ for all $R$ in $N$.

(ii) If $y$ is in $L_x$, there exists an $R$ in $N$ such that $R(x) = y$.

Consequently, theorems (3.2) and (3.3) (the proof of these theorems is taken from Zorn's work) can be combined into the following:

"A transformation $F$ in $N$ with the fixed point $p$ is either a rotation or is nilpotent."

That this theorem is true under the hypotheses and definitions of section 2 is shown in section 4.

Now (4.4) might have been written: "If $F$ is in $N$, either $|F(x)| = |x|$ for all $x$, or $|F(x)| < |x|$ for all $x$ other than $p$."

In order to interpret Schwarz' lemma in this form, Zorn imposes further restrictions on $S$ and $N$:

(i) $S$ is connected and contains more than one point;

(ii) $S$ is locally connected;

(iii) $S$ has no cut points;

(iv) if $q \neq p$, $S - L_q$ is not connected;

(v) $S$ is not representable as a finite sum of interiors of spheres.
Requirement (iv) permits a definition of the relation "$L_x$ is interior to $L_y$." This in turn permits the following definitions:

$|x| = |y|$ if $x$ and $y$ lie in the same circumference, and $|x| < |y|$ if $L_x$ is interior to $L_y$.

It may be noted that (iv), interpreted in the sense of the present paper, is a direct consequence of \[2.2.1\]. On the other hand, (iv), in the sense of Zorn's work, is a consequence of \[2.3\] and \[4.4\]. When we consider that Zorn needs (iv) in order to interpret his version of Schwarz's lemma metrically, it becomes intuitively more apparent what role \[2.3\] plays in the proofs of \[4.2\] and \[4.3\].

The writer wishes to thank Dr. R. F. Beekenbach, who suggested the topic of this thesis, and Drs. H. E. Bray and S. Mandelbrojt, who gave many valuable criticisms and suggestions during the preparation of the paper.

The Rice Institute,
June, 1941.