Re-visiting the Proofs of the Geometric Forms of the Hahn Banach Theorem Concerning the Separation of Convex Sets

B. G. Akuchu
Department of Mathematics
University of Nigeria
Nsukka, Nigeria

K. T. Nwigbo
Department of Mathematics
University of Nigeria
Nsukka, Nigeria

This article is distributed under the Creative Commons by-nc-nd Attribution License. Copyright © 2021 Hikari Ltd.

Abstract
We show that the proofs of the first and second geometric forms of the Hahn Banach Theorem, as presented by several authors (see e.g [1], [2]), appear to have some obscurities. We present a cutting-edge proof of the theorems void of any ambiguities.

Introduction and main results
The Hahn Banach Theorems (see e.g [1], [2]) are amongst the corner stone theorems in Mathematical Analysis (Functional Analysis, to be more precise). There are basically two forms (Analytic and Geometric) and several other consequent theorems. Many authors (see e.g [1-8]) have studied these theorems. The analytic form and its several consequences deal with the extension of linear functionals from subspaces to the whole space, in linear and normed linear spaces. It guarantees the existence of bounded linear functionals in non-empty non-degenerate normed linear spaces.
The geometric forms, whose proofs are our main concern, are theorems which provide conditions under which convex sets are separated in normed linear spaces. Several authors (see e.g [1], [2]) have studied these theorems. Before we give the statements of the theorems, we recall the following definitions

**Definition 1 (see e.g [1]):** Let $X$ be a linear space and $A \subset X$, $B \subset X$. We say the hyperplane of equation $[f = \alpha]$ separates $A$ and $B$ in:(i) a general sense if $f(x) \leq \alpha, \forall x \in A$ and $f(x) \geq \alpha, \forall x \in B$ (ii) a strict sense if there exist $\epsilon > 0$ such that $f(x) \leq \alpha - \epsilon, \forall x \in A$ and $f(x) \geq \alpha + \epsilon \forall x \in B$ (in other words, $f(x) < \alpha \forall x \in A$ and $f(x) > \alpha \forall x \in B$).

We now give the statements of the theorems, with a view to study the proofs of the geometric forms, as presented by several authors.

**Theorem 1 (Hahn Banach, Analytic Form) (see e.g [1]):** Let $X$ be a linear space and $M$ be a proper subspace of $X$. Let $p : X \to \mathbb{R}$ (reals) be a sublinear functional and $f : M \to \mathbb{R}$ be a linear functional from $M$ into $\mathbb{R}$, satisfying $f(x) \leq p(x), \forall x \in M$. Then

(i) There exists a linear functional $F : X \to \mathbb{R}$ which is an extension of $f$.

(ii) $F(x) \leq p(x) \forall x \in X$.

**Theorem 2 (Hahn Banach, Consequence of Analytic Form) (see e.g [1]):** Let $X$ be a normed linear space and $M$ be a proper subspace of $X$. Let $f : M \to \mathbb{R}$ be a bounded linear functional on $M$. Then

(i) There exists a bounded linear functional $F$ on $X$, which extends $f$.

(ii) $\|F\| = \|f\|$.

**Theorem 3 (Hahn Banach, First Geometric Form) (see e.g [1], [2]):** Let $X$ be a normed linear space. Let $A \subset X$, $B \subset X$ be two convex, non-empty and disjoint sets. Suppose that $A$ is open. Then there exists a closed hyperplane that separates $A$ and $B$ in the general sense.

**Theorem 4 (Hahn Banach, Second Geometric Form) (see e.g [1], [2]):** Let $X$ be a normed linear space. Let $A \subset X$, $B \subset X$ be two convex, non-empty and disjoint sets. Suppose that $A$ is closed and $B$ is compact. Then there exists a closed hyperplane that separates $A$ and $B$ in the strict sense.

The method of proving theorem 3 (First Geometric Form) is to invoke theorem 1 (Analytic Form). Therefore, the hypotheses of theorem 1 are created. That is, we seek a subspace $M$ of $X$, a sublinear functional, $p$, on $X$ and a linear functional $f$ on $M$. In a bid to get the sublinear functional, $p$, on $X$, several authors (see e.g [1], [2]) made the following definition:

**Definition 2:** Let $X$ be a normed linear space. Let $K \subset X$ be an open,
convex set with \(0 \in K\). Define the function \(p : X \to \mathbb{R}\) by

\[
p(x) = \inf\{\alpha > 0 : \alpha^{-1}x \in K\},
\]

for all \(x \in X\). Then \(p\) is called the gauge of \(K\). With this definition, the authors (see e.g [1]) stated and proved the following lemma:

**Lemma 1** (see e.g [1] Lemma 3.23, [2] Lemma 1.2): The gauge of \(K\) satisfies the following conditions:

\[
\begin{align*}
G_1 & : \quad p(\lambda x) = \lambda p(x), \quad \forall \ x \in X, \ \lambda > 0 \\
G_2 & : \quad \text{There exists a constant } k > 0 \text{ such that } 0 \leq p(x) \leq k||x|| \text{ for all } x \in X \\
G_3 & : \quad K = \{x \in X : p(x) < 1\} \\
G_4 & : \quad p(x + y) \leq p(x) + p(y) \text{ for all } x, y \in X.
\end{align*}
\]

**Remark 1:** The proof of \(G_3\) (see e.g [1]) is heavily dependent on the fact that \(0 \in K\). Furthermore, the proof of \(G_4\) is dependent on \(G_3\) (and hence on the fact that \(0 \in K\)).

To complete the proof of theorem 3 (Hahn Banach, First Geometric Form), the authors (see e.g [1], [2]) stated and proved the following lemma:

**Lemma 2** (see e.g [1] Lemma 3.25, [2] Lemma 1.3): Let \(K \subset X\) be an open, convex and nonempty subset of \(X\) and let \(x_0 \in X, \ x_0 \notin K\). Then there exists \(F \in X^*\) such that \(F(x) < F(x_0)\), for all \(x \in K\) (In particular, the hyperplane of equation \([F = F(x_0)]\) separates \(\{x_0\}\) and \(K\) in the general sense).

In proving this lemma, authors (see e.g [1], [2]) assume that \(0 \in K\) by translation, so that as in lemma 1, the gauge of \(K\) is the sublinear functional \(p : X \to \mathbb{R}\). The authors (see e.g [1], [2]) now generated the subspace \(M := \{x = \lambda x_0 : \lambda \in \mathbb{R}\}\) and defined a functional \(f : M \to \mathbb{R}\) by \(f(x) = f(\lambda x_0) := \lambda\). It is easily proven that \(p\) and \(f\) satisfy the hypotheses of the sublinear functional and the linear functional in theorem 1 (Hahn Banach, Analytic Form), respectively. Hence by theorem 1 (Hahn Banach, Analytic Form), there exists an extension \(F : X \to \mathbb{R}\), satisfying the required conditions.

Finally, to complete the proof of theorem 3 (Hahn Banach, First Geometric Form), authors (see e.g [1], [2]) employ the technique of lemma 2 by setting \(K = A - B := \{a - b : a \in A, \ b \in B\}\) and \(\{x_0\} = \{0\}\).

The authors now claim (the proofs are routine exercises) that:

(i) \(0 \notin K\) (which is obvious)

(ii) \(K\) is open

(iii) \(K\) is convex

They now invoke lemma 2 and use the fact that \(F(x_0) = F(0) = 0\) to show and conclude that \(F\) separates \(A\) and \(B\) in the general sense.
In this paper, we point out that the technique of setting \( K = A - B \), \( \{x_0\} = \{0\} \) and invoking lemma 2 directly to get the conclusions appear to have some obscurities. We make the following observations:

(i) If \( \{x_0\} = \{0\} \), then the subspace \( M \) as generated in the proof of lemma 2 (see [1] lemma 3.25, [2] Lemma 1.3) above is \( M = 0 \).

(ii) The first condition that \( K = A - B \) satisfies is that \( 0 \notin K \) (which is true, otherwise \( A \) and \( B \) are not disjoint).

We are of the opinion that if \( 0 \notin K \), then \( K \) does not possess a gauge (see lemma 1). To further buttress the fact that \( K \) is not translated so that \( 0 \in K \), we see from the concluding part of the proof of Theorem 1.6 of [2] that "...By lemma 1.3 there exists \( f \in X^* \) such that \( f(z) < f(0) = 0 \) \( \forall z \in K \)."

Otherwise, we have an absurdity. We want to make the following two-items emphasis:

a) If \( 0 \notin K \), then (i)-(iv) cannot hold for \( K \) and hence \( K \) does not have a gauge function. Consequently, we do not have a sublinear functional.

b) If \( K \) is translated so that \( 0 \in K \), then from the concluding part of the theorem 1.6 of [2] (which is our theorem) which says that "...By lemma 1.3 there exists \( f \in X^* \) such that \( f(z) < f(0) = 0 \) \( \forall z \in K \)."

These are apparently obscurities in the proof of the theorem.

We re-define the conditions satisfied by the gauge of a subset \( K \) of the normed linear space \( X \) and use these to prove the above theorem, using a different approach.

**Lemma 3:** Let \( X \) and \( K \) be as in definition 2. Then the gauge of \( K \) satisfies the following conditions:

\[ G_1 : \quad p(\lambda x) = \lambda p(x), \quad \forall x \in X, \quad \lambda > 0 \]

\[ G_2 : \quad \text{There exists a constant} \quad k > 0 \quad \text{such that} \quad 0 \leq p(x) \leq k ||x|| \quad \text{for all} \quad x \in X \]

\[ G_3 : \quad K = \{ x \in X : p(x) \leq 1 \} \]

\[ G_4 : \quad p(x + y) \leq p(x) + p(y) \quad \text{for all} \quad x, y \in X. \]

The difference between Lemma 1 and lemma 3 is the definition of \( K \) in \( G_3 \). The proof of \( G_3 \) in lemma 3 follows from the fact that, if \( \epsilon \in (0, 1) \) is chosen arbitrary and \( x \in K \), then \( (1 + \epsilon)x \in K \). This implies \( p(x) \leq \frac{1}{(1+\epsilon)} \). Letting \( \epsilon \to 0 \), we have \( p(x) \leq 1 \). Conversely, if \( p(x) \leq 1 \), then there exists \( \alpha \in (0, 1] \) such that \( \alpha^{-1}x \in K \). Now, we have \( 0 \in K \), \( \alpha^{-1}x \in K \). By the convexity of \( K \) and since \( \alpha \in (0, 1] \), we have \( \alpha(\alpha^{-1}x) + (1 - \alpha)0 \in K \). Thus \( K = \{ x \in X : p(x) \leq 1 \} \). It is worth noting that condition \( G_3 \) is very crucial in the proof of condition \( G_4 \) whose proof follows as in [1], except that \( p(x) \leq 1 \) is employed throughout.

With this, we now give a cutting-edge proof of theorem 3.

**Cutting-Edge Proof of Theorem 3.** Set \( K = A - B := \{ a - b : a \in A, \ b \in B \} \). Just as above, \( K \) is non-empty, open and convex. Choose \( x_0 \in X \) arbitrary such that \( x_0 \notin K \). Let \( y_0 \in K \) and set \( x_0^* = x_0 - y_0 \) and \( K^* = K - y_0 \). Obvi-
Re-visiting the proofs of geometric forms

Thus, $K^*$ has a guage $p$. Let $M = \{ x : x = \lambda x_0^*, \lambda \in R{\text{reals}} \}$ and define $f : M \rightarrow R{\text{ (reals)}}$ by $f(x) = f(\lambda x_0^* ) \equiv \lambda$. It is easily seen that $M$ is a subspace of $X$ and $f$ is a linear functional on $M$. Obviously $p(0) = f(0) = 0$ (for $0 = \lambda$,). Also, $p(x) = p(\lambda x_0^*) = \lambda p(x_0^*) \geq \lambda = f(x)$ (for $0 < \lambda$, since $x_0^* \notin K^*$). Next, for $\lambda < 0$, we have $f(x) = f(\lambda x_0^*) = \lambda < 0 \leq p(x)$. Hence $f(x) \leq p(x), \forall x \in M$.

Therefore, from the analytic form of the Hahn Banach theorem, $f$ has a continuous linear extension on $X$, say $F$, satisfying $F(x) \leq p(x), \forall x \in X$. In particular, for any $x \in K^*$, $F(x) \leq p(x) \leq 1 = f(x_0^*) = F(x_0^*)$. Now, recall that $x \in K^*$ implies $x = y - y_0$, for some $y \in K$, and $x_0^* = x_0 - y_0$. So $F(y) - F(y_0) = F(y - y_0) = F(x) \leq F(x_0^*) = F(x_0 - y_0) = F(x_0) - F(y_0)$ and thus $F(y) \leq F(x_0)$ for all $y \in K$. Recall that any $y \in K$ is of the form $a - b$, where $a \in A$ and $b \in B$. Since in particular, $0 = x_0 \notin K$, we have $F(a) - F(b) = F(a - b) = F(y) \leq F(x_0) = F(0) = 0$. This completes the proof.

**Remark 2**: Our proof is straightforward and is void of any ambiguities. Furthermore, the proof of lemma 2 is rather a corollary of our proof.

**Remark 3**: Observe that with $K$ defined as in $G_3$ in lemma 1, one can only get strict separation, as in the following theorem. This means we would loose separation in the general sense and note that separation in the general sense implies separation in the strict sense. The converse is however not true.

**Theorem 5**: Let $X$ be a normed linear space. Let $A \subset X, B \subset X$ be two convex, non-empty and disjoint sets. Suppose that $A$ is open. Then there exists a closed hyperplane that separates $A$ and $B$ in the strict sense.

**Proof**: Let $p$ be as in definition 2, satisfying the conditions of lemma 1. The proof of conditions $G_1 - G_4$ follows as in [1], with the exception that in proving $G_3$, we choose a FIXED $\alpha \in (0, 1)$. This yields $K = \{ x \in X : p(x) < 1 \}$, as in [2].

Set $K = A - B := \{ a - b : a \in A, b \in B \}$. Just as above, $K$ is non-empty, open and convex. Choose $x_0 \in X$ arbitrary such that $x_0 \notin K$. Let $y_0 \in K$ and set $x_0^* = x_0 - y_0$ and $K^* = K - y_0$. Obviously, $K^*$ is non-empty, open and convex. Also, $x_0^* \notin K^*$ and $0 = y_0 - y_0 \in K^*$. Thus, $K^*$ has a guage $p$. Let $M = \{ x : x = \lambda x_0^*, \lambda \in R\{\text{reals}\} \}$ and define $f : M \rightarrow R\{\text{reals}\}$ by $f(x) = f(\lambda x_0^* ) \equiv \lambda$. It is easily seen that $M$ is a subspace of $X$ and $f$ is a linear functional on $M$. Obviously $p(0) = f(0) = 0$ (for $0 = \lambda$,). Also, $p(x) = p(\lambda x_0^* ) = \lambda p(x_0^*) \geq \lambda = f(x)$ (for $0 < \lambda$, since $x_0^* \notin K^*$). Next, for $\lambda < 0$, we have $f(x) = f(\lambda x_0^*) = \lambda < 0 \leq p(x)$. Hence $f(x) \leq p(x), \forall x \in M$. 

ousse, $K^*$ is non-empty, open and convex. Also, $x_0^* \notin K^*$ and $0 = y_0 - y_0 \in K^*$. Thus, $K^*$ has a guage $p$. Let $M = \{ x : x = \lambda x_0^*, \lambda \in R\{\text{reals}\} \}$ and define $f : M \rightarrow R\{\text{reals}\}$ by $f(x) = f(\lambda x_0^* ) \equiv \lambda$. It is easily seen that $M$ is a subspace of $X$ and $f$ is a linear functional on $M$. Obviously $p(0) = f(0) = 0$ (for $0 = \lambda$,). Also, $p(x) = p(\lambda x_0^* ) = \lambda p(x_0^*) \geq \lambda = f(x)$ (for $0 < \lambda$, since $x_0^* \notin K^*$). Next, for $\lambda < 0$, we have $f(x) = f(\lambda x_0^*) = \lambda < 0 \leq p(x)$. Hence $f(x) \leq p(x), \forall x \in M$. 

ousse, $K^*$ is non-empty, open and convex. Also, $x_0^* \notin K^*$ and $0 = y_0 - y_0 \in K^*$. Thus, $K^*$ has a guage $p$. Let $M = \{ x : x = \lambda x_0^*, \lambda \in R\{\text{reals}\} \}$ and define $f : M \rightarrow R\{\text{reals}\}$ by $f(x) = f(\lambda x_0^* ) \equiv \lambda$. It is easily seen that $M$ is a subspace of $X$ and $f$ is a linear functional on $M$. Obviously $p(0) = f(0) = 0$ (for $0 = \lambda$,). Also, $p(x) = p(\lambda x_0^* ) = \lambda p(x_0^*) \geq \lambda = f(x)$ (for $0 < \lambda$, since $x_0^* \notin K^*$). Next, for $\lambda < 0$, we have $f(x) = f(\lambda x_0^*) = \lambda < 0 \leq p(x)$. Hence $f(x) \leq p(x), \forall x \in M$.
Therefore, from the analytic form of the Hahn Banach theorem, \( f \) has a continuous linear extension on \( X \), say \( F \), satisfying \( F(x) \leq p(x) \), \( \forall x \in X \). In particular, for any \( x \in K^\ast \), \( F(x) \leq p(x) < 1 = f(x^\ast_0) = F(x^\ast_0) \). Now, recall that \( x \in K^\ast \) implies \( x = y - y_0 \), for some \( y \in K \), and \( x^\ast_0 = x_0 - y_0 \). So \( F(y) - F(y_0) = F(y - y_0) = F(x) < F(x^\ast_0) = F(x_0 - y_0) = F(x_0) - F(y_0) \) and thus \( F(y) < F(x_0) \) for all \( y \in K \). Recall that any \( y \in K \) is of the form \( a - b \), where \( a \in A \) and \( b \in B \). Since in particular, \( 0 = x_0 \notin K \), we have \( F(a) - F(b) = F(a - b) = F(y) < F(x_0) = F(0) = 0 \). This completes the proof.

**Remark 4:** This result concerns the strict separation of convex sets under different conditions. The conditions on the subsets \( A \) and \( B \) in this theorem are different from those in theorem 4, which require that \( A \) be closed and \( B \) be compact.

**Competing Interests:** We declare that there are no competing interests surrounding this article.

**References**


Received: July 17, 2021; Published: August 19, 2021