ON THE HAHN-BANACH THEOREM

LAWRENCE NARICI

Dedicated to my friend Jean Schmets on the occasion of his retirement

Abstract. I love the Hahn-Banach theorem. I love it the way I love Casablanca and the Fontana di Trevi. It is something not so much to be read as fondled. What is “the Hahn-Banach theorem?” Let \( f \) be a continuous linear functional defined on a subspace \( M \) of a normed space \( X \). Take as the Hahn-Banach theorem the property that \( f \) can be extended to a continuous linear functional on \( X \) without changing its norm. Innocent enough, but the ramifications of the theorem pervade functional analysis and other disciplines (even thermodynamics!) as well. Where did it come from? Were Hahn and Banach the discoverers? The axiom of choice implies it, but what about the converse? Is Hahn-Banach equivalent to the axiom of choice? (No.) Are Hahn-Banach extensions ever unique? They are in more cases than you might think, when the unit ball of the dual is “round,” as for \( \ell_p \) with \( 1 < p < \infty \), for example, but not for \( \ell_1 \) or \( \ell_\infty \). Instead of a linear functional, suppose we substitute a normed space \( Y \) for the scalar field and consider a continuous linear map \( A: M \to Y \). Can \( A \) be continuously extended to \( X \) with the same norm? Well, sometimes. Unsurprisingly, it depends on \( Y \), more specifically, on the “geometry” of \( Y \): If the unit ball of \( Y \) is a “cube,” as for \( Y = (\mathbb{R}^n, \|\cdot\|_\infty) \) or \( Y = \text{real } \ell_\infty \), for example, then for any subspace \( M \) of any \( X \), any bounded linear map \( A: M \to Y \) can be extended to \( X \) with the same norm. This is not true if \( Y = (\mathbb{R}^n, \|\cdot\|_p), n > 1 \), for \( 1 < p < \infty \), despite the topologies being identical. The cubic nature of the unit ball does not suffice, however—if \( Y = c_0 \), the extendibility dies. This article traces the evolution of the analytic form as well as subsequent developments up to 2004.

1. What is it?

The two principal versions of the Hahn-Banach theorem are as a continuous extension theorem (analytic form) and as a separation theorem (geometric form) about separating convex sets by means of a continuous linear functional that takes different values on the sets.

Analytic Forms.

Dominated version. Let \( f \) be a continuous linear functional defined on a subspace \( M \) of a real vector space (no norm) \( X \), \( p \) a sublinear functional defined on \( X \) and \( f \leq p \) on \( M \); \( f \) can be extended to a linear functional \( F \) defined on \( X \) with \( F \leq p \).

\[
\begin{align*}
F & : X \\
\downarrow & \\
F \leq p & \\
\hline
f & : M \\
\rightarrow & \mathbb{R} \\
f & \leq p
\end{align*}
\]

For complex spaces, we mainly need some absolute values: If $X$ is complex, and $p$ a seminorm such that $|f| \leq p$ on $M$ then $|F| \leq p$.

**Norm-preserving version.** If $X$ is normed space over $K = \mathbb{R}$ or $\mathbb{C}$ and $f : M \to K$ is a continuous linear functional then there exists a continuous linear functional $F$ extending $f$ defined on all of $X$ such that $|F| = |f|$.

**Geometric Form.** Let $X$ be a real or complex topological vector space. In any real or complex TVS $X$, if the linear variety $x + M$ does not meet the open convex set $G$ then there exists a closed hyperplane $H$ containing $x + M$ that does not meet $G$ either. Mazur 1933 deduced the geometric form from the analytic form; he made no mention of the converse possibility. In a 1941 article, Dieudonné [1981b] refers to the geometric form as the Hahn-Banach theorem, so he was apparently aware of the equivalence of the two. It is first called the geometric form by Bourbaki.

**The analytic form is a cousin of Tietze’s theorem that a bounded continuous $f : K \to [a, b]$ defined on a closed subset $K$ of a normal space $T$ possesses a continuous extension $F : T \to [a, b]$ with the same bounds. The geometric form resembles Urysohn’s lemma about separating disjoint closed subsets of a normal space by a continuous function. Urysohn’s lemma is usually proved by induction, the geometric Hahn-Banach theorem by transfinite induction.**

In the interest of keeping size reasonable, I consider only the analytic form in the sequel. There are many—denumerably, I suspect—other versions of the theorem, for vector lattices, modules, boolean algebras, bilinear functionals, groups, semigroups and more. It has many applications not only outside functional analysis but outside mathematics. Feinberg and Lavine [1983], for example, develop thermodynamics using the Hahn-Banach theorem, Neumann and Velasco [1994] apply Hahn-Banach type theorems to develop feasibility results on the existence of flows and potentials and Delbaen and Schachermayer [1994] use it to develop a fundamental theorem of asset pricing.

### 2. The Obvious Solution

Suppose that $X$ is just a vector space—no norm—over $K = \mathbb{R}$ or $\mathbb{C}$ and the linear functional $f$ maps a subspace $M$ into $K$. An easy way to extend $f$ to $X$ is to take an algebraic complement $N$ of $M$, consider the projection $P_M$ on $M$ along $N$ and take $F = f \circ P_M$. In effect, take $F$ to be 0 outside $M$. Will this technique work for continuous linear functionals $f$ defined on a closed (extend $f$ by continuity to $M$ if $M$ is not closed) subspace $M$ of a topological vector space $X$? If $P_M$ is continuous, then $F = f \circ P_M$ is a continuous linear extension of $f$ though not necessarily of the same norm (cf. Sec. 7.1). Generally, however, we cannot rely on this method because $P_M$ is continuous if and only if $N$ is a topological complement of $M$ [Narici and Beckenstein 1985, (5.8.1)(a)] and uncomplemented subspaces are common—$c_0$, for example, is an uncomplemented subspace of $\ell_\infty$, so there is no continuous projection of $\ell_\infty$ onto $c_0$ [Narici and Beckenstein 1985, Ex. 5.8.1, $C [0, 1], L_p [0, 1]$ and $L_p, 1 \leq p \leq \infty, p \neq 2$, have closed uncomplemented subspaces [Köthe 1969, pp. 430–1] and for $0 < p < 1$, no finite-dimensional subspace of $L_p [0, 1]$, has a topological complement [Köthe 1969, p. 158]. In fact, any Banach space $X$ has uncomplemented closed subspaces unless $X$ is linearly homeomorphic to a Hilbert space [Lindenstrauss and Tzafriri 1971]. Some instances in which a subspace $M$ of a locally convex space is complemented are $M$ finite-dimensional or codimensional,
or $M$ a closed subspace of a Hilbert space, in which case its orthogonal complement $M^\perp$ is a topological complement. We say a little more about the Hilbert space situation in Sec. 7.2, this being a case in which $f \circ P_M$ is the only continuous linear extension of $f$ with the same norm.

3. The Times

Throughout the nineteenth and early twentieth centuries the function concept was significantly broadened, analysis became “geometrized,” the idea of structure emerged and the standards of rigor improved greatly; the techniques of Euclidean geometry became the standard. Functional analysis evolved from the desire to do analysis on function spaces, treating functions the way points in $\mathbb{R}$ or $\mathbb{R}^2$ had been dealt with.

3.1. The Evolution of the Function Concept; Analysis on Spaces of Functions. Fourier series played an important role not only in the enlargement of the function concept, but in analysis in general throughout the 19th century. Dirichlet boldly proffered the characteristic function of the rationals, the Dirichlet function, in 1829 as a function for which the integrals for the coefficients of its Fourier series “lose every meaning.” There was no disagreement about the lack of integrability, but was this a function? something which could not be graphed? At the time, with no formal definition of function—indeed, with practically no formal definitions of anything—“function” had only its intuitive meaning. It was tacitly required that it had to be graphable and it meant essentially “elementary function:” polynomials, trigonometric functions, exponentials and logarithms. In 1837 Dirichlet proposed that any correspondence between the points of interval $[a,b]$ and points of $\mathbb{R}$ be considered a function. In view of the strange behavior of even continuous functions such as Riemann’s continuous-for-irrational $x$, discontinuous-for-rational-$x$ in 1854 and Weierstrass’s nowhere differentiable continuous function in 1874, it became apparent that more latitude was clearly necessary. The ability to uniformly approximate “strange” functions by trigonometric functions helped render them acceptable as did Weierstrass’s 1885 demonstration that any continuous function could be uniformly approximated by polynomials.

Consideration of functions whose domains were other than subsets of $\mathbb{R}$ or $\mathbb{C}$ has a venerable history. Circa 100 B.C. Zenodorus considered the isoperimetric problem—among the closed plane curves of a given length, find the one that encloses the most area—so, even at this early date there was consideration of numbers associated with curves and choosing the curve or curves corresponding to the least number. A similar situation occurred when the Bernoulli brothers considered the brachistochrone problem—from the class of curves connecting two points, associate a number, a time, with each and choose a curve corresponding to the least number. Throughout the 1700’s it was common to associate numbers with curves by means of definite integrals. Not only was mapping curves into numbers common, since the early 1800’s so were function-to-function mappings such as differential operators, Laplace transforms, and shift operators. In spite of this long history of linking functions with numbers or other functions, it took until the late 1800’s to formalize the notion of a function as a correspondence between elements of arbitrary sets. [“Set” or “class” of functions had been in common use in the early 19th century, well before Boole in 1847 or Cantor. Volterra 1887 spoke of numerical- and $\mathbb{R}^n$-valued functions defined on the set of all continuous curves (linee or lignes) in a
square and then did something decisively different: he proposed doing analysis on them—limits, continuity, derivatives. This was possible because there were already several distinct ways of judging proximity of functions that arose from the various notions of convergence of a sequence of functions that developed in the latter half of the 19th century. Volterra called these new kinds of functions \textit{funzioni dipenditi da linee} or \textit{fonctions de ligne} where by \textit{ligne} he meant a continuous image of \([0,1]\) in the unit square. Peano realized early the exotic possibilities of such a broad criterion for “curve.” In part to make his point, he invented his space-filling curve. Undaunted by this cautionary example, Hadamard pursued it. In 1902 he wrote a short note \cite{Bull. Soc. Math. France 30, 40–43} on Volterra’s derivatives of \textit{fonctions de lignes}. In 1903 \cite{Comptes Rendus 136, 351–354} he abandoned \textit{fonction de ligne} and called the new functions of functions \textit{fonctionnelles}, analysis on them \textit{analyse fonctionnelle} and gave our subject its name. Hadamard’s student Paul Lévy wrote a book, \textit{Leçons d’analyse fonctionnelle}, in 1922 in which he divided \textit{calcul fonctionnelle} into \textit{algèbre fonctionnelle} and \textit{analyse fonctionnelle}; the \textit{algèbre} dealt with problems whose unknowns were ordinary functions, the \textit{analyse} with problems in which the unknowns were \textit{fonctionnelles}.

3.2. Structure and Isomorphism. Throughout the 19th century, the idea that concretely different things could be the same in some crucial sense gestated, i.e., the notion of isomorphism. In what appears to be the first use of the term, the German chemist Eilhard Mitscherlich formulated the \textit{principle of isomorphism} of crystals in 1819, concerning similarity of geometric forms of crystal structures and the chemical consequences of such an arrangement. Unlike our usage of \textit{isomorphic}, chemists say \textit{isomorphous} as in sodium nitrate and calcium sulfate have \textit{isomorphous} crystal structure.

Contemporaneously, geometers such as Gauss, Lobachevsky and Bolyai—and Klein’s Erlanger program—created non-Euclidean geometries and reformulated classical geometry. These developments influenced the idea of “space.” Hilbert’s 1899 \textit{Grundlagen der Geometrie}, and its many subsequent editions, launched the abstract mathematics of the 20th century. The qualitative leap that Hilbert and others made is that they did not try to define points and lines and planes as Euclid had attempted; rather, they accepted these notions as “atoms,” without intrinsic content. Not only didn’t you know what they were, you couldn’t. In a 1941 letter to Frege, Hilbert wrote:

\begin{quote}
If among my points I consider some systems of things (e.g., the system of love, law, chimney sweeps . . . .) and then accept only my complete axioms as the relationships between these things, my theorems (e.g., the Pythagorean) are valid for these things also.
\end{quote}

Through the medium of the axiomatic system, mathematics was about to attain a level of abstraction hitherto unknown.

The idea of a \textit{vector} had been around in the 19th century but it meant \textit{n}-tuple. With infinite-dimensional vector spaces in mind, in Chapter IX of his 1888 book, Peano gave a rather modern axiomatic definition of \textit{vector space} and \textit{linear map}. Pincherle had already been writing about \textit{spazio funzionale}, \textit{operazioni funzionali}, \textit{calcolo funzionale} and linear operators on complex sequence spaces, so he was quite receptive to Peano’s ideas. Pincherle wrote a book about vector spaces (\textit{Le operazioni distributive e le loro applicazioni all’analisi} with U. Amaldi) in 1901, but
the idea was mostly ignored. Even Hilbert in his research on $\ell_2$ was indifferent to its vector space structure. It was not until Riesz, Helly, Hahn, Banach and Wiener endowed vector spaces with a norm that interest in them ignited. More generally, the idea of \textit{structure} had arrived. \textit{Group} (a term coined by Galois) was defined in 1895, \textit{field} in 1903. A comprehensive framework for the various notions of limit and continuity for particular sets of functions that Volterra and Hadamard were investigating materialized with M. Fréchet’s abstract metric space in 1904, ideas refined in his 1906 thesis. Fréchet used Jordan’s term, \textit{écart}, for metric distance. Although Fréchet liberally used spatial imagery, he did not coin the term \textit{metric space}; Hausdorff (probably) was the first to do so in his \textit{Mengenlehre} [1914].

3.3. \textbf{Spatial Imagery and the Euclidean Renaissance.} By the late nineteenth and early twentieth centuries, suggestive geometric terminology and spatial imagery were commonly applied to arbitrary sets of ‘points.’ Hilbert and his school spoke of \textit{orthogonal expansions}; in 1913, Riesz described the solution of systems of homogeneous equations

$$f_i(x) = a_{i1}x_1 + \cdots + a_{in}x_n = 0, \quad 1 \leq i \leq n$$

as an attempt to find $x = (x_1, \ldots, x_n)$ \textit{orthogonal} to the linear span $[f_1, \ldots, f_n]$ where $f_i = (a_{i1}, \ldots, a_{in})$, i.e., solving the equations is an attempt to identify the orthogonal complement of $[f_1, \ldots, f_n]$. Significantly, the ‘equations,’ the $f_i$, achieved vector status; they were of the same species as the ‘variables.’ Peano’s \textit{Calcolo Geometrico} in 1888 and Minkowski’s \textit{Geometrie der Zahlen} in 1896 are two seminal works in the geometrizing of analysis. Minkowski tinctured analysis with ideas about \textit{convexity} in $\mathbb{R}^n$. (Even so, in a 1941 article on the Hahn-Banach theorem mentioned earlier, Dieudonné 1981b defines \textit{convexity}, so it wasn’t standard even by 1941.) Minkowski defined support hyperplane, support function and proved the existence of a support hyperplane at every boundary point of a convex body. Helly extended Minkowski’s notions about convexity from $\mathbb{R}^n$ to normed sequence spaces (see Sec. 4.4).

Contemporaneously, Euclidean methodology became established: theorem–proof arguments to make deductions from explicitly stated assumptions, the type of rigor that became prevalent in the 20th and early 21st centuries. The standards were raised so much that most earlier work looks shabby by comparison. For example, from the 17th into the 19th century, infinite series were usually treated the same as polynomials with no regard for convergence.

4. \textbf{Origins}

Attempts to solve infinite systems of linear equations led to early versions of the Hahn-Banach theorem as well as to the creation of the general normed space. The analog of the diagonalizability conditions for finite systems of linear equations, the necessary and sufficient condition for solvability of an infinite system of linear equations, is compatibility between the linear equations and the scalars which can be described as the continuity of a certain linear functional. (As normed spaces had not been defined yet, this was not the interpretation given at the time.) A key figure is F. Riesz (no surprise) who proved variants of the Hahn-Banach theorem for $L_p[0,1]$, $\ell_p$, and $BV[a,b]$. In the course of his investigations with $C[a,b]$, Riesz very nearly defined the general normed space. Helly also proved special cases of
the Hahn-Banach theorem, defined a general normed sequence space and a dual of a sequence space. Hahn, Banach and Wiener subsequently defined the general real normed space. Hahn and Banach each independently proved the Hahn-Banach theorem for real normed spaces.

4.1. Systems of Linear Equations. Two prominent problems of the late nineteenth century were:

Moment problems.
Given a sequence \((c_n)\) of numbers, find a function \(x\) with those ‘moments,’ i.e., such that

\[
\int_0^1 t^n x(t) \, dt = c_n \quad \text{for every } n \in \mathbb{N}
\]

Fourier series.
Given a sequence \((g_n)\) of cosines or sines and \((c_n)\) of numbers, find a function \(x\) for which the \(c_n\) are the Fourier coefficients, i.e., such that

\[
\int_{-\pi}^{\pi} x(t) g_n(t) \, dt = c_n \quad \text{for every } n \in \mathbb{N}
\]

We can rephrase these in a more general setting. Let \(X\) be a normed space with dual \(X'\), let \(S\) be a set, and let \(\{c_s : s \in S\}\) be a collection of scalars.

\textbf{(V) The vector problem.}
Let \(\{f_s : s \in S\}\) be a collection of bounded linear functionals on \(X\). Find \(x \in X\) such that \(f_s(x) = c_s\) for every \(s\).

and its dual:

\textbf{(F) The functional problem.}
Let \(\{x_s : s \in S\}\) be a collection of vectors from \(X\). Find \(f \in X'\) such that \(f(x_s) = c_s\) for every \(s\).

If \(X\) is reflexive then solving (F) also solves (V), for given “vectors” \(\{f_s : s \in S\} \subset X'\) there exists \(h \in X''\) such that \(h(f_s) = c_s\) for every \(s\). Now choose \(x \in X\) such that \(h(f_s) = f_s(x)\) for every \(s\).

Suppose \(X\) is a reflexive normed space and consider a simple vector problem: Given functionals \(f, g \in X'\) and scalars \(a\) and \(b\), find \(x \in X\) such that \(f(x) = a\) and \(g(x) = b\). If \(f\) and \(g\) are linearly independent, for any scalars \(c\) and \(d\), take \(h((cx + dy)) = ca + db\). Then extend the continuous linear functional \(h\) to \(H \in X''\) by the Hahn-Banach theorem. Finally, choose \(x \in X\) such that for all \(\phi \in X'\), \(H(\phi) = \phi(x)\) to solve the problem. If \(f\) and \(g\) are linearly dependent, then the scalars have to be “compatible.” If \(g = 2f\), say, then we must have \(b = 2a\). More generally, given functionals \(f_1, f_2, \ldots, f_n\) and scalars \(c_1, c_2, \ldots, c_n\), if \(\sum a_i f_i = 0\) then we must also have \(\sum a_i c_i = 0\). This type of compatibility is guaranteed by conditions (*) and (***) in Sec. 4.2 and (***) of Theorem 1 where it can be viewed as a continuity condition.

4.2. Riesz. Motivated by Hilbert’s work on \(L_2[0, 1]\), Riesz 1910 invented the spaces \(L_p[0, 1], 1 < p < \infty\); in 1913 he considered the \(\ell_p\) spaces. He generalized the moment and Fourier series problems [1910, 1911] to the vector problem (LP) below. In solving (LP), he inadvertently solved a functional problem and created an early Hahn-Banach extension theorem. Riesz solved the vector problem in the reflexive spaces \(L_p[0, 1], p > 1\). Simultaneously, he solved an associated functional problem in \(L_p' = L_q\) which yielded a special case of the Hahn-Banach theorem.
(LP) Let $S$ be a set. For $p > 1$ and $1/p + 1/q = 1$, given $y_s$ in $L_q[a,b]$ [equivalently, consider the functionals $f_s$ in Eq. (1)] and scalars $c_s$ for each $s \in S$, find $x$ in $L_p[a,b]$ such that

$$f_s(x) = \int_a^b x(t)y_s(t) \, dt = c_s \quad \text{for every } s$$

For there to be such an $x$, he showed that the following necessary and sufficient connection between the $y$'s and the $c$'s had to prevail: There exists $K > 0$ such that for any finite set of indices $s$ and scalars $a_s$

$$\left| \sum a_sc_s \right| \leq K \left( \int_a^b \left| \sum a_sy_s \right|^q \right)^{1/q} = K \left\| \sum a_sy_s \right\|_q$$

in today’s language. Condition (\*) means that if the $y$'s are linearly dependent, if $\sum a_sy_s = 0$ for some finite set of scalars $a_s$, then $\sum a_sc_s = 0$ as well. Thus, if we consider the linear functional $g$ on the linear span $M = \{y_s : s \in S\}$ of the $y$'s in $L_q[a,b]$ defined by taking $g(y_s) = c_s$, the $g$ so obtained is well-defined. Not only that, for any $y$ in $M$, $|g(y)| \leq K \|y\|_q$ on $M$, so $g$ is continuous on $M$. If there is an $x$ in $L_p$ which solves (LP), then $g$ has a continuous extension $G$ to $L_q$, namely, for any $y$ in $L_q$,

$$\langle x, y \rangle = G(y) = \int x(t)y(t) \, dt \quad [G(y_s) = c_s (s \in S)]$$

Thus, Riesz showed that (LP) is solvable if and only if a certain linear functional $g$ defined on a subspace of $L_q$ is continuous; if it is solvable, $g$ is also special in that it is a restriction of a continuous linear functional defined on all of $L_q$.

In a paper written and submitted in 1916 but not published until 1918, Riesz turned to the following vector problem.

(BV) Given $y_s \in C[a,b]$, and scalars $c_s (s \in S)$, find $x \in BV[a,b]$

(functions of bounded variation on $[a,b]$) such that

$$f_s(x) = \int_a^b y_s(t) \, dx(t) = c_s \quad (s \in S)$$

He solved it with a necessary and sufficient condition—continuity, again—very much like (\*), namely: There exists $K > 0$ such that for any finite set of indices $s$ and any scalars $a_s$

$$\left| \sum a_sc_s \right| \leq K \sup_{s \in [a,b]} \left| \sum a_sy_s \right| = K \left\| \sum a_sy_s \right\|_\infty$$

in modern notation.

4.3. The First Normed Space. In his 1918 article, Riesz used “norm” and the notation $\|x\| = \sup |x|[0,1]$ for $x$ in the funktionraum $C[a,b]$. He observed that $\|x\|$ is generally positive, and is zero only when $x(t)$ vanishes identically. Furthermore . . . (for any scalar $a$ and any $x, y \in C[a,b]$)

$$|ax(t)| = |a| \|x\|, \quad [\text{and}] \quad \|x + y\| \leq \|x\| + \|y\|$$

By the distance of $x, y$ we understand the norm $\|x - y\| = \|y - x\|$.
He didn’t define the general normed space but he came mighty close. Although he was working in particular spaces, Riesz intuited that these were only galaxies in a greater universe. Even in 1913, in the introduction to his book, Les systèmes d’équations linéaires à une infinité d’inconnues, he said:

Strictly speaking, our study is not part of the theory of functions but rather can be considered as a first stage of a theory of functions of infinitely many variables.

The notation $\|\cdot\|$ was first used by Hilbert’s student E. Schmidt 1908, as in “Gram-Schmidt,” who called $x = (a_n) \in \ell_2$ “normiert” if $\sum |a_n|^2 < \infty$. When he defined norm in his book, Banach used $|\cdot|$, not $\|\cdot\|$.

Riesz unintentionally extended some continuous linear functionals in solving $(LP)$ and $(BV)$. Helly made a qualitative leap: he went directly to the extension.

4.4. Enter Helly. Despite the brevity of his mathematical career—only five journal publications—the Austrian mathematician Eduard Helly (1884–1943) made significant contributions to functional analysis. Dieudonné 1981a [p. 130] characterized Helly’s 1921 article as “a landmark in the history of functional analysis.” For an excellent discussion of Helly and his work, see Butzer et al. 1980 and 1984. By means different from and simpler than Riesz’s, Helly also solved $(BV)$ in 1912 and proved special cases of the Hahn-Banach and Banach-Steinhaus theorems in $C[a,b]$. A bullet through the lungs in September 1915—a wound that ultimately caused his death—ended his stint at the eastern front as a soldier in the Austrian army in World War I. He spent almost the next five years as a prisoner of war, enduring eastern Siberia’s frigidity since 1916. He did not return to Vienna until mid-November of 1920 by way of Japan, the Middle East and Egypt. I would like to think that the opportunity for distraction afforded by thinking about mathematics helped sustain him during that awful period. In any case he revisited $(BV)$ in 1921 with a perspective that definitively anticipates the Hahn-Banach theorem. Probably in the belief that all spaces were reflexive, Helly tried to solve $(BV)$ by means of a corresponding functional problem, then finding the “vector” (in $X''$, actually) that corresponded to the functional. He defined a general normed sequence space and a dual. Some highlights of his approach are:

- **Normed sequence spaces** Helly dealt with a general vector subspace $X$ of the space $C^N$ of complex sequences equipped with a norm, an *abstandfunktion* or “distance function.” He did not use the notation $\|\cdot\|$. This is general enough to cover the $\ell_p$ spaces and many others such as $L_2$ which can be identified with $\ell_2$. Helly realized that a norm generalized what Minkowski (in $R^n$) called the gauge of a convex body.

- **Dual spaces** Given a normed subspace $X$ of $C^N$, Helly took as its “dual space”

$$X' = \left\{ (u_n) \in C^N : \sum_{n \in \mathbb{N}} x_n u_n < \infty \quad \text{for all } (x_n) \in X \right\}$$

i.e., $(u_n)$ such that $(x_n u_n)$ is summable for all $(x_n) \in X$. If $X = c$ or $c_0$, then $X' = \ell_1$ by this method; if $X = \ell_1$, then $X' = \ell_\infty$. If $X = \ell_\infty$, the $X'$ you get this way is only part of what we call the dual of $X$ today.

- **Norms the dual space** For $x = (x_n) \in X$ and $u = (u_n) \in X'$, Helly defines a bilinear form $\langle \cdot, \cdot \rangle$ on $X \times X'$ (that makes $(X,X')$ a dual pair) by taking for $x$ in $X$ and $u$ in $X'$ $\langle x, u \rangle = \sum_{n \in \mathbb{N}} x_n u_n$. Using an idea of Minkowski’s, he considers a
dual norm for $X'$ by taking

$$
\|u\| = \sup \left\{ \frac{|\langle x, u \rangle|}{\|x\|} : x \neq 0 \right\}
$$

The dual norm on $X$ obtained from $X'$ by this technique yields the original norm on $X$. (Nowadays such pairs $(X, X')$, subject to absolute convergence of $\sum x_n u_n$, are called Köthe sequence spaces and $\alpha$-duals, respectively.) The dual norm also permitted Helly to consider successive duals. He sought to solve the following vector problem:

\[ (SQ) \text{ Given sequences } f_n = (f_{nj}) \text{ from } X' \subset C^N, \text{ and a sequence } (c_n) \in C^N, \text{ find } x \in X \text{ such that } \langle x, f_n \rangle = \sum_{j \in N} x_j f_{nj} = c_n \text{ for each } n \in N \]

He tried to solve (SQ) by finding $h \in X''$ such that $h(f_n) = c_n$ and using reflexivity. He discovered that the $x \in X$ corresponding to $h$ did not always exist, thus showing that some spaces are not reflexive. As part of his solution, with a condition like (*) above, Helly proved a restricted version of Theorem 1, below, by extending a bounded linear functional $f$ from a subspace $M$ to the whole space.

The key step was, for $x$ not in $M$, to find a linear $F$ such that

$$
F : M \oplus \mathbb{C}x \rightarrow K \quad F \leq k \|\cdot\|
$$

the same idea that Hahn 1927 and Banach 1929 used to prove the Hahn-Banach theorem.

4.5. Hahn and Banach. Riesz 1918 had considered $C[a,b]$ as a normed space. Hahn 1922, Wiener 1922 and Banach 1923 took the final step: each independently defined the general real normed space. Considering Hahn and Banach’s awareness of what Helly did (complex normed sequence spaces), it is surprising that that neither mentions complex scalars—indeed, Banach does not consider them even in his 1932 book. Wiener 1923 observed that complex scalars could be used just as well as real scalars. Hahn called a norm a Massbestimmung. Hahn and Banach each required completeness. Banach removed it in his book, distinguishing between normed and Banach spaces. Each considered the general problem of extending a continuous linear functional defined on a subspace of a general normed space, not a sequence space, as Helly did. Hahn published the norm-preserving form of the theorem in 1927; Banach proved it independently and published his result in 1929. He mentioned Helly, acknowledged Hahn’s priority in his book 1932, and generalized it to the dominated version. Although he made no further use of the greater generality, the more general result was useful with the advent of locally convex spaces. Each of them used Helly’s technique to obtain the theorem—reduce the problem to the case of a one-vector extension—but instead of ordinary induction, they used transfinite induction. (The Zorn’s lemma equivalent of transfinite induction commonly used today did not arrive until 1935.) Hahn also formally introduced the notion of dual space (polare Raum), noted that $X$ is embedded in its second dual $X''$ and defined reflexivity (regularité).

Generalizing Riesz’s results for $L_p[0,1]$ and $BV[a,b]$, Banach 1932 extended a result of Helly 1912 to obtain the compatibility condition (***) of Theorem 1
Theorem 1. (Helly-Banach) Let $X$ be a real (or complex) normed space, let $\{x_s\}$ and $\{c_s\}$, $s \in S$, be sets of vectors and scalars, respectively. Then there is a continuous linear functional $f$ on $X$ such that $f(x_s) = c_s$ for each $s \in S$ if and only if there exists $K > 0$ such that for all finite subsets $\{s_1, \ldots, s_n\}$ of $S$ and scalars $a_1, \ldots, a_n$

\[ \left| \sum_{i=1}^{n} a_i c_{s_i} \right| \leq K \left\| \sum_{i=1}^{n} a_i x_{s_i} \right\|. \]

Banach used the Hahn-Banach theorem to prove Theorem 1 but Theorem 1 implies the Hahn-Banach theorem: Assuming that Theorem 1 holds, let $\{x_s\}$ be the vectors of a subspace $M$, let $f$ be a continuous linear functional on $M$; for each $s \in S$, let $c_s = f(x_s)$. Since $f$ is continuous, condition (***)) is satisfied and $f$ possesses a continuous extension to $X$.

5. The complex case

F. Murray 1936 discovered the intimate relationship between the real and complex parts of a complex linear functional $f$, namely that $\text{Re} \, f(ix) = \text{Im} \, f(x)$.

He reduced the complex case to the real case, and proved the complex version of the dominated version of the Hahn-Banach theorem for subspaces of $L_p(a,b]$ for $p > 1$. Murray’s perfectly general method was used and acknowledged by Bohnenblust and Sobczyk 1938 who proved it for arbitrary complex normed spaces. They were the first to call it the Hahn-Banach theorem. Also by reduction to the real case, Soukhomlinov 1938 and Ono 1953 obtained the theorem for vector spaces over the complex numbers and the quaternions.

Hustad 1973, Holbrook 1975 and Mira 1982 present unified approaches. Instead of reduction to the real case, each utilizes an intersection property of $\mathbb{R}$, $\mathbb{C}$, and the quaternions to prove it for all three simultaneously. We discuss these intersection properties in Sec. 6.1.

6. A Theorem for Linear Maps? (not Functionals)

Notation. Except for Secs. 10 and 11, in the sequel $X$ and $Y$ denote at least normed spaces over $K = \mathbb{R}$ or $\mathbb{C}$; for any vector $x$ and $r > 0$, $B(x,r) = \{y : \|y - x\| \leq r\}$.

Can we replace the scalar field $K = \mathbb{R}$ or $\mathbb{C}$ in the Hahn-Banach theorem by a normed space $Y$? If $A : M \to Y$ is a bounded linear map on the closed subspace $M$ of $X$, is there a linear extension $\bar{A} : X \to Y$ of $A$ such that $\|\bar{A}\| = \|A\|$? If such an $\bar{A}$ exists for any such $A$ on any subspace $M$ of any normed space $X$, we will say that $Y$ is extendible.

As we shall see, in the real case we can let $Y$ be $\ell_\infty(n)$ for any $n$, even $n = \infty$—a space whose unit ball is a “cube”—but not $c_0$, even though its unit ball is also cubic. The extendible spaces are characterized by intersection properties of their closed balls (Theorems 2 and 3) which resemble compactness. The first one proved and the easiest to state (for the real case) is the binary intersection property, namely:
If $\mathcal{B}$ is a collection of mutually intersecting closed balls, then $\bigcap \mathcal{B} \neq \emptyset$.

As an external characterization, a space is extendible if and only if it is norm-isomorphic to some $(C(T, K), \|\|_\infty)$ where $T$ is an extremally disconnected compact Hausdorff space.

6.1. Intersection Properties. If $Y$ is extendible then we must be able to continuously extend the identity map $1: Y \to Y$ to $\overline{Y}$ on the completion $\overline{Y}$ of $Y$. Thus if the sequence $(y_n)$ from $Y$ converges to $y \in \overline{Y}$, then $\overline{1} y_n = y_n \to y = \overline{1} y \in Y$. Extendible spaces are therefore complete.

Since $c_0$ is uncomplemented in $\ell_\infty$ [Narici and Beckenstein 1985, p. 87], there is no continuous projection of $\ell_\infty$ onto $c_0$. This means that the identity map $y \mapsto y$ of $c_0$ onto $c_0$ does not have a continuous extension to $\ell_\infty$ so $c_0$ is not extendible. More generally, we see that an extendible space must be complemented in any Banach space in which it is norm-embedded.

As there is no loss of generality in doing so, we assume that $\|A\| \leq 1$ on $\text{min}$ the following discussion. The key step to extension is the ability to extend a function $A$ from $M$ to $\overline{A}$ defined on $M \oplus Kx$ for any $x \notin M$. To preserve the bound, it is necessary and sufficient to choose a value $y$ for $\overline{Ax}$ that satisfies $\|\overline{Ax} - Am\| = \|y - Am\| \leq \|x - m\|$ for all $m \in M$. Thus, a permissible value $y$ must lie in $B(\overline{Am}, \|x - m\|)$ for every $m \in M$, must be in $\bigcap_{m \in M} B(\overline{Am}, \|x - m\|)$, in other words. Since $\|y\| = \|Am\| \leq \|m\|$, we need to know what spaces $Y$ satisfy

$$(IP) \quad \bigcap_{m \in M} B(y, \|x - m\|) \neq \emptyset \quad \text{for any } y \in B(0, \|m\|) \text{ and any } x \notin M$$

By the Hahn-Banach theorem, we know that $IP$ is satisfied in $\mathbb{R}$ and $C$ no matter what $M$ and $X$ are, but the presence of the $M \subseteq X$ here is troublesome—it has nothing to do with $Y$. For the sake of determining purely internal characterizations of $IP$ (Theorems 2 and 3), consider the following intersection properties.

Let $S$ be a set and $\mathcal{B} = \{B(y_s, r_s) : y_s \in Y, r_s > 0, s \in S\}$ be a collection of closed balls in $Y$. If $\bigcap \mathcal{B} \neq \emptyset$ whenever:

- $\{B(y_s, r_s) : y_s \in Y, r_s > 0, s \in S\}$ is mutually intersecting then $Y$ has the binary intersection property;
- $\bigcap \{B(f(y_s), r_s) : s \in S\} \neq \emptyset$ (in $K$) for any $f$ in the unit ball of $Y'$ then $Y$ has the weak intersection property;
- for any $B(y_k, r_k) \in \mathcal{B}$, $k = 1, 2, \ldots, n$, $n \in \mathbb{N}$ and $b_1, b_2, \ldots, b_n \in K$, $\sum_{k=1}^n b_k = 0$ implies that $\|\sum_{k=1}^n b_k y_k\| \leq \sum_{k=1}^n |b_k| r_k$ then $Y$ has Holbrook’s intersection property.

**Theorem 2.** General Case A Banach space $Y$ over $K = \mathbb{R}$ or $C$ is extendible if and only if

(a) [Goodner 1976] If the Banach space $X$ contains $Y$, there is a continuous projection $P$ of norm 1 of $X$ onto $Y$ (Y is “1-complemented” in $X$);

(b) The identity map $1: Y \to Y$ can be extended to a linear map of the same norm to any Banach space $X$ containing $Y$.

(c) $Y$ is topologically complemented in each space in which it is norm-embedded.

(d) [Hustad 1973] $Y$ has the weak intersection property.

(e) [Holbrook 1975, Mira 1982] $Y$ satisfies Holbrook’s intersection property.

(f) [Hasumi 1958] There exists an extremally disconnected [open sets have open closures] compact Hausdorff space $T$ such that $Y$ is norm-isomorphic to
By the Banach-Stone theorem, $T$ is unique up to homeomorphism.

Mira 1982 corrected an error in Holbrook’s argument and also showed that if $K = \mathbb{C}$, the previous conditions are equivalent to Holbrook’s intersection property being satisfied for all sets of three elements; if $K = \mathbb{R}$ or a non-Archimedean valued field (assuming that $X$ has a norm which also satisfies the ultrametric triangle inequality) then we need only require it for sets of two elements, while for $K = \mathbb{Q}$ the quaternions, it suffices that it be satisfied for sets of five.

Nachbin 1950 and Goodner 1950 each showed that a real extendible space whose unit ball had an extreme point was linearly isometric to some $C(T, \mathbb{R})$ where $T$ is an extremally disconnected compact Hausdorff space. Nachbin had conjectured that the extreme point hypothesis was redundant. Kelley 1952 and Goodner 1976 validated his conjecture. For an extremally disconnected compact Hausdorff space $T$, the function $e(t) \equiv 1$ is an extreme point of the unit ball of $(C(T, \mathbb{R}), \|\cdot\|_\infty)$; hence, by Theorem 2(f), a necessary condition for a space to be extendible is that its unit ball possess extreme points.

Akilov 1948 provides another necessary condition for finite-dimensional spaces [cf. Goodner 1950, Cor. 4.8]: If a real finite-dimensional space $Y$ is smooth, then it is not extendible. [A point $u$ of the surface $S$ of the unit ball of $Y$ is a smooth point if there is a unique supporting hyperplane at $u$ which is equivalent to Gateaux differentiability of the norm at $u$. $Y$ is called smooth, if it is smooth at each $u \in S$.]

**Theorem 3.** **REAL CASE** A real Banach space $Y$ is extendible if and only if any of the conditions of Theorem 2 are satisfied as well as if and only if:

(a) [Nachbin 1950, Goodner 1950, Kelley 1952] $Y$ has the binary intersection property.

In (b) and (c), the balls can be enlarged somewhat.

(b) [Lindenstrauss 1964] Any family $B = \{B(y_s, r_s) : y_s \in Y, s \in S\}$ of mutually intersecting closed balls is such that for every $r > 0, \bigcap_{s \in S} B(y_s, (1 + r) r_s) \neq \emptyset$.

(c) [Davis 1977] Any family $B = \{B(y_s, 1) : y_s \in Y, s \in S\}$ of mutually intersecting closed unit balls is such that for every $r > 0, \bigcap_{s \in S} B(y_s, 1 + r) \neq \emptyset$.

(d) [Nachbin 1950, Goodner 1950, 1976, Kelley 1952] $Y$ is norm-isomorphic to a complete Archimedean ordered vector lattice with order unit.

In addition to the sources cited, see also Secs. 8.8 and 10.5 of Narici and Beckenstein 1985 and Herrero 2003, pp. 149f.

**Theorem 4.** **REFLEXIVITY** [Goodner 1950, Theorem 6.8; Nachbin 1950, Theorem 5] A real extendible space is reflexive if and only if it is finite-dimensional.

For real separable spaces, we therefore have:

**Theorem 5.** **SEPARABLE SPACES** Let $Y$ be a real extendible normed space. Then $Y$ is separable if and only if

(a) [Goodner 1960] $Y$ is reflexive.

(b) [Goodner 1950] $Y$ is finite-dimensional.

(c) [Goodner 1960] There exists a finite discrete space $T$ such that $X$ is norm-isomorphic to $(C(T, \mathbb{R}), \|\cdot\|_\infty)$. 
We need to define a few terms to state the next characterizations (Theorem 6) of extendibility of real spaces.

**Definition.** Let $B$ be a bounded subset of a normed space $X$.

(a) The *diameter* $d(B)$ of $B$ is $\sup \{\|x - y\| : x, y \in B\}$.

(b) The *radius* $r(B)$ of $B$ is $\inf \{r > 0 : B \subset B(y, r), y \in X\}$.

In addition to boundedness, suppose that $B$ is closed and convex for (c) and (d).

(c) $B$ is *diametrically maximal* if for every $x \notin B$, $d(\{x\} \cup B) > d(B)$.

(d) $B$ has *constant width* $d > 0$ if for each $f \in X'$ with $\|f\| = 1$, $\sup f(B - B) = d$.

Sets of constant width must be diametrically maximal; the two notions coincide in any two-dimensional space as well as in $n$-dimensional spaces with the Euclidean norm [Eggleston 1965]. They are distinct in certain three-dimensional spaces. It follows from Franchetti 1977, Moreno 2005 and Moreno, et. al. 2005 that if $Y = (C(T, \mathbb{R}), \|\cdot\|_\infty)$, where $T$ is a compact Hausdorff space, they coincide if and only if $T$ is extremally disconnected; this yields Theorem 6(b).

**Theorem 6.** Radial descriptions, real spaces $A$ real Banach space $Y$ is extendible if and only if any of the conditions of Theorem 3 are satisfied as well as if and only if:

(a) [Davis 1977] For every bounded subset $B$ of $Y$, the diameter $d(B) = 2r(B)$.

(b) For every closed bounded convex subset $B$ of $Y$, $B$ has constant width if and only if $B$ is diametrically maximal.

6.2. Examples on Extendible Spaces.

- As $(\mathbb{R}, |\cdot|)$ is a complete, Archimedean ordered vector lattice with order unit, it is extendible.

- When Helly 1912 proved the fundamental lemma—the one-dimensional extension—to his version of the Hahn-Banach theorem, he observed that a family of mutually intersecting closed intervals $\{[a_s, b_s] : s \in S\}$ of $\mathbb{R}$ has nonempty intersection, i.e., that $\mathbb{R}$ has the binary intersection property. He generalized this [1923] to his intersection theorem, namely that a family $\{B_s : s \in S\}$ of compact convex subsets of $\mathbb{R}^n$ has nonempty intersection if any $n + 1$ of them meet; he generalized it to a topological theorem in 1930.

- $(\mathbb{R}^2, \|\cdot\|_2)$ is not extendible because it does not have the binary intersection property: There clearly exist three mutually intersecting circles whose intersection is empty. For essentially the same reason, none of $(\mathbb{R}^n, \|\cdot\|_p)$, $1 < p < \infty$, are extendible for $n > 1$; one could also argue that they are not extendible because their unit balls are smooth.

- [Nachbin 1950, Theorem 3] The only real normed spaces of finite dimension $n$ that are extendible are those that are norm-isomorphic to $\ell_\infty(n)$. Since the map $e_1 \mapsto e_1 + e_2$, $e_1 \mapsto e_1 - e_2$, defined on the standard basis vectors $e_1$ and $e_2$ of $\mathbb{R}^2$, is a linear isometry of real $\ell_1(2)$ onto real $\ell_\infty(2)$, it follows that real $\ell_1(2)$ is extendible.

- Real Hilbert spaces of dimension $\geq 2$ do not have the binary intersection property, so are not extendible.

- The real space $(B(T, \mathbb{R}), \|\cdot\|_\infty)$ of bounded real-valued functions on any set $T$ has the binary intersection property. If $T = \{1, 2, \ldots, n\}$ or $\mathbb{N}$ then $B(T, \mathbb{R}) = (\mathbb{R}^n, \|\cdot\|_\infty) = \text{real } \ell_\infty(n)$ or real $\ell_\infty$, respectively.
Since \((C([0,1],\mathbb{R}),\|\cdot\|_\infty)\) is uncomplemented in \((B([0,1],\mathbb{R}),\|\cdot\|_\infty)\), \(C[0,1]\) is not extendible.

Extendibility is a geometric rather than a topological property.

The topologies of real \(\ell_p(2)\), are the same for all \(1 \leq p \leq \infty\) but only \(\ell_1(2)\) and \(\ell_\infty(2)\) are extendible.

6.3. **The domain.** What normed spaces \(X\) have the property that any continuous linear map \(A\) of any subspace \(M\) into any normed space \(Y\) has a linear extension with same norm?

\[
\bar{A} : X \quad \text{fixed} \quad \|\bar{A}\| = \|A\|
\]

\[
A : M \quad \longrightarrow \quad Y \quad A,M,Y \quad \text{arbitrary}
\]

If \(X\) is a Hilbert space and \(P_M\) the orthogonal projection on the closed subspace \(M\), then \(\bar{A} = A \circ P_M\) is a norm-preserving extension of \(A\) to \(X\). And this is just about the only situation in which there are norm-preserving extensions. Kakutani 1939 (real case) and Bohnenblust 1942 and Sobczyk 1944 in the complex case (cf. also Saccoman 2001) showed that the only Banach spaces \(X\) with this property are Hilbert spaces and those \(X\) of dimension \(\leq 2\).

6.4. **Superspaces and Functionals.** Suppose \(M\) is a normed space over \(K\) where \(K\) is \(\mathbb{R}, \mathbb{C}\), the quaternions or even a non-Archimedean valued field. Suppose further that for any \(X\) in which \(M\) is norm-embedded that every bounded linear functional on \(M\) has an extension of the same norm to \(X\). With one proof that works for all four fields, Mira 1982 showed that these \(M\) are just the extendible spaces.

6.5. **Superspaces and Linear Maps.** Let \(M\) and \(Y\) be real Banach spaces and \(A : M \rightarrow Y\) is a continuous linear map. Suppose the real Banach space \(X\) contains \(M\) as a closed subspace. Suppose further that \(\bar{A}\) is a linear extension of \(A\) of the same norm to \(X\). \(M\) is such that such extensions exist for all \(X, Y\) and \(A\) if and only if \(M\) is extendible [Nachbin 1950].

\[
\bar{A} : X \quad \|\bar{A}\| = \|A\|
\]

\[
A : M \quad \longrightarrow \quad Y \quad M \quad \text{fixed}
\]

It follows from Theorem 2(a), that there must be a continuous projection \(P\) of norm 1 of \(X\) onto \(M\). Gajek et al. 1995 characterize such \(M\) by means of properties of the Gateaux derivative of the norm on \(X\); for just continuous extensions, as opposed to norm-preserving ones, it suffices to be able to renorm \(X\) so that there is a continuous projection \(P\) of norm 1 of \(X\) onto \(M\). See also Ostrovskii 2001 and Chalmers et al. 2003.

7. **Uniqueness of the Extension**

Uniqueness of continuous extensions of continuous linear functionals is closely linked to smoothness of the normed space \(X\). For example, if \(M\) is a subspace of the normed space \(X\) and \(f \in M'\) attains its norm at a smooth point, then \(f\) has a unique extension of the same norm to \(X\). More generally, unique extensions are guaranteed for any continuous linear functional on any subspace of \(X\) if and only if \(X'\) is strictly convex (= strictly normed = rotund) which implies that \(X\) is smooth.
7.1. **Non-Uniqueness.** Consider the subspace \( M = \mathbb{R} \subset (\mathbb{R}^2, \| \cdot \|_2) \) and the linear functional \( f(a, 0) = a \) defined on \( M \). Let \( y \in \mathbb{R}^2 \) be a unit vector of angle \( \beta \neq 0, \pi \) with the \( x \)-axis. The subspace \( N \) spanned by \( y \) is a topological complement of \( M \). For any such \( N \), the projection \( P_M \) on \( M \) along \( N \) is a continuous extension of \( f \) and \( f \circ P_M \) is a continuous extension of \( f \) of norm \( |\text{csc} \beta| \), so there are infinitely many continuous extensions of \( f \) but only one of the same norm (\( \| f \| = 1 \)), namely when \( \beta = \pm \pi/2 \) (when \( N = M^\perp \)); this, incidentally, yields the extension of \( f \) of minimal norm, the smallest value of \( |\text{csc} \beta| \). Distinct extensions of \( f \) of the same norm, 1, are given by \( F(a, b) = a + b \) and \( G(a, b) = a - b \).

For an instance in which there are infinitely many extensions of a continuous linear functional of the same norm, consider the subspace \( M \) of constant functions of the Banach space \( (C[0, 1], \| \cdot \|_\infty) \) of complex-valued continuous functions on \([0, 1]\) and the continuous linear functional \( f: M \to \mathbb{C}, x \mapsto x(0) \). Clearly \( \| f \| = 1 \). For any \( t \in [0, 1] \), the evaluation map \( F_t: C[0, 1] \to \mathbb{C}, x \mapsto x(t) \) extends \( f \) and is of norm 1.

7.2. **Unique Extensions of the Same Norm: Special Cases.** Preservation of the norm clearly limits the choices that can be made for an extension and there are always extensions that do not preserve the norm.

**Theorem 7.** If \( f \) is a continuous linear functional defined on the closed proper subspace \( M \) of the normed space \( X \) over \( K = \mathbb{R} \) or \( \mathbb{C} \) then there are continuous linear extensions \( F \) of \( f \) with \( \| F \| > \| f \| \).

**Proof.** Choose a unit vector \( u \not\in M \) and let \( d \) be the positive distance from \( u \) to \( M \). Then, for any \( a \in K \) and any \( m \in M, |a|d \leq \| au - am \| \) indeed, 

\[
(*) \quad |a|d \leq \| au + m \| \quad (a \in K, m \in M)
\]

Choose a scalar \( b > \| f \| \). Define \( F \) on \( M \oplus Ku \) by taking \( F(m + au) = f(m) + ab \) for any \( a \in K \) and \( m \in M \). To see that \( F \) is continuous, suppose \((m_n)\) and \((a_n)\) are sequences from \( M \) and \( K \), respectively, such that \( m_n + a_n u \to 0 \). By (*) for every \( n, |a_n|d \leq \| a_nu + m_n \| \to 0 \); hence \( a_n \to 0 \) and \( m_n \to 0 \). Therefore \( F(a_nu + m_n) = a_n b + f(m_n) \to 0 \). Since \( F(u) = b > \| f \| \), it follows that \( \| F \| > \| f \| \).

As we argue next, extensions of the same norm are unique in Hilbert spaces.

Suppose \( f \) is a bounded linear functional defined on a closed subspace \( M \) of a Hilbert space \( (X, \langle \cdot, \cdot \rangle) \), and let \( P_M \) be the orthogonal projection on \( M \). Extend \( f \) to \( F = f \circ P_M \). Since orthogonal projections are continuous, so is \( F \). Therefore, by the Riesz representation theorem, there exist unique \( m \in M \) and \( n \in M^\perp \) \( (X = M \oplus M^\perp) \) such that \( F(\cdot) = \langle \cdot, m + n \rangle \) and \( \| F \| = \| m + n \| \). Since \( 0 = F(n) = \| n \|^2 \), it follows that \( n = 0 \). Hence \( \| F \| = \| m \| = \| f \| \). If \( G \) is any extension of \( f \) of the same norm, a similar argument shows that \( G(\cdot) = \langle \cdot, m \rangle = F(\cdot) \).

The situation for certain subspaces of the \( \ell_p \) spaces, \( 1 < p < \infty \), is similar. Let \( \{e_n\} \) be the standard basis for \( \ell_p \), let \( S \) be a subset of \( N \) and let \( f \) be a bounded linear functional defined on the closed linear span \( M \) of \( \{e_n : n \in S\} \). For \( q = p/(p-1), \| f \|^q = \sum_{n \in S} |f(e_n)|^q \). Given an extension \( F(a_n) = \sum_{n \in S} a_n f(e_n) + \sum_{n \notin S} a_n F(e_n) \) of \( f \) of the same norm, then \( \| f \|^q = \| F \|^q \) implies that \( F(e_n) = 0 \) for all \( n \notin S \).

These uniqueness results for Hilbert and \( \ell_p \) spaces are special cases of the Taylor-Foguel theorem of Sec. 7.5.
7.3. Uniqueness of Dominated Extensions. Suppose \( f \) is a linear functional defined on a subspace \( M \) of a real vector space \( X \) and that \( f \) is dominated by a sublinear functional \( p \) (defined on \( X \)) on \( M \). To prove that \( f \) can be extended from \( M \) to \( F \) defined on \( M \oplus \mathbb{R}x_0 \), \( x_0 \notin M \), in such a way \( F \) is still dominated by \( p \), a number \( c \) is chosen arbitrarily between \( a = \sup \{-p(-m-x_0) - f(m): m \in M\} \) and (the larger quantity) \( b = \inf \{p(m+x_0) - f(m): m \in M\} \) as the value for \( F(x_0) \). Herein lies the non-uniqueness of \( F \). If these two values are equal for every \( x \), i.e., if \( p \), \( f \) and \( M \) are such that, for each \( x \in X \),

\[
(2) \quad \sup \{-p(-m-x) - f(m): m \in M\} = \inf \{p(m+x) - f(m): m \in M\}
\]

then there is only one extension \( F \) of \( f \) with \( F \leq p \); conversely, if \( F \) is unique, then Eq. (2) must hold [de Guzmán 1966; cf. Herrero 2003, Th. 5.2.1]. The assertion of Eq. (2) is equivalent to, for each \( x \notin M \),

\[
\sup \{f(m) - p(m-x): m \in M\} = \inf \{p(m+x) - f(m): m \in M\}
\]

Bandyopadhyay and Roy 2003 characterize when a single linear functional dominated by a sublinear functional \( p \) on a subspace \( M \) of a real vector space has a unique extension to the whole space dominated by \( p \) in terms of nested sequences of “\( p \)-balls” in a quotient space; by considering the canonical embedding of \( M \) in its bidual \( M'' \), they characterize unique extendibility of elements of \( M' \) in terms of sequences from \( M \).

If \( X \) is complex [cf. Hererro 2003, Cor. 5.2.6], \( p \) a seminorm defined on \( X \) and \(|f| \leq p \) on the subspace \( M \) then \( f \) has a unique extension \( F \) to \( X \), \(|F| \leq p \), if and only if for every \( x \in X \),

\[
\sup \{-p(-x-m) - \text{Re} (f(m)): m \in M\} = \inf \{p(x+m) - \text{Re} f(m): m \in M\}
\]

7.4. Unique Extensions for Points and Subspaces—Best Approximations from \( M^0 \).

**Notation.** \( M \) and \( U \) (or \( U(X) \)) denote, respectively, a closed subspace and the unit ball of a Banach space \( X \) in this section and we consider only norm-preserving extensions of \( f \in M' \) (the continuous dual of \( M \)) to an element \( F \in X' \). \( L(X) \) and \( K(X) \) denote, respectively, the spaces of all continuous linear operators and compact operators of \( X \) into \( X \).

**One Point.** If \( f \in M' \) attains its norm at a smooth point, then \( f \) has a unique extension to \( X \) [Holmes 1975, p. 176].

**One subspace—Phelps’s theorem.** The seminal result characterizing subspaces \( M \) for which elements of \( M' \) have unique extensions is that of Phelps 1960: Continuous linear forms \( f \) on \( M \) have unique extensions to \( f \in X' \) if and only if the annihilator \( M^0 = \{ u \in X' : u |_M = 0 \} \) is Cébysev in \( X' \), in other words each \( g \in X' \) has a unique best approximation \( h \in M^0 \), an \( h \) such that \( ||g-h|| = \inf \{ ||g-u|| : u \in M^0 \} \).

We next consider a sufficient condition on a subspace for it to have unique extensions.
HB-subspaces. M is called an HB-subspace if there is a projection P on X' whose kernel is M⁰ and, for each f ∈ X' such that f ≠ Pf, ∥Pf∥ < ∥f∥ and ∥f - Pf∥ ≤ ∥f∥. Hennefeld 1979 showed that if M is an HB-subspace, then each f ∈ M' is uniquely extendible to X. Oja 1984 showed that a subspace may have unique extensions but not be an HB-subspace—for X = R² normed by ∥(a, b)∥ = max(|a|, |a + b|/2), the subspace R has unique extensions but is not an HB-subspace. Oja 1997 gets some equivalent conditions for M to be an HB-subspace and also shows that X is an HB-subspace of its bidual whenever K(X) is an HB-subspace of L(X).

A subclass of the HB-subspaces for which it is easier to give examples is the M-ideals.

M-ideals. M is called an M-ideal if there is a projection P on X' whose kernel is M⁰ and, for each f ∈ X' such that f ≠ Pf, ∥f∥ = ∥f∥ + ∥f - Pf∥.

Examples of M-ideals.
(a) If X is a B*-algebra, then any closed 2-sided ideal in X is an M-ideal [Smith and Ward 1978].
(b) Thus, spaces (C(T), || ||∞) of continuous functions on a compact set T are well supplied with M-ideals: For any t ∈ T, the maximal ideal Mt = {x ∈ C(T) : x(t) = 0} is an M-ideal.
(c) If T is a locally compact and Hausdorff, the M-ideals of (C∞(T), || ||∞), the continuous scalar-valued functions that vanish at infinity, are precisely MF = {x ∈ C∞(T) : x(F) = {0}} where F is closed in T [Behrends 1979, p. 40].
(d) For X = L₂(µ), for some measure µ, the subspace K(X) of compact operators is an M-ideal in L(X).
(e) The space c₀ of null sequences is an M-ideal in ℓ∞. Thus, any continuous linear functional on c₀ has a unique Hahn-Banach extension to ℓ∞ (Harmand et al. 1993, Proposition 1.12).

M-ideals and intersection properties. M-ideals may be characterized internally in various ways by intersection properties of balls. For real X, the closed subspace M is an M-ideal if and only if M satisfies the 3-ball property, namely that if three open balls B₁, B₂, B₃ have nonempty intersection and each meets M, then M ∩ (∁₁=³ Bᵢ) ≠ ∅ [Alfsen and Effros 1972; cf. Behrends 1979, p. 46f.]. Behrends proved it for subspaces of complex spaces in 1991. Oja and Poldvere 1999 consider a related condition called the “2-ball sequence property” and show that M satisfies the 2-ball sequence property if and only if each f ∈ M' has a unique extension to F ∈ X'.

Costara and Popa 2001 give further examples of subspaces for which Hahn-Banach extensions are unique.

7.5. Unique Extensions for all Subspaces—Rotund Dual. If the surface SU of the unit ball of X contains no nontrivial line segments (i.e., SU consists entirely of extreme points), then X is called strictly convex or rotund. Taylor 1939 proved that if the dual X' is strictly convex then any f ∈ M' has a unique extension to F ∈ X' of the same norm. He proved the converse when X is reflexive. Foguel 1958 removed the reflexivity, thereby showing that the normed spaces X for which each continuous linear functional on any subspace of X has a unique linear extension of the same norm are those X with strictly convex dual. Strict convexity of X' implies that X (not X') is smooth. Since Hilbert spaces and ℓ₂, 1 < p < ∞, have
strictly convex duals, bounded linear functionals on subspaces of either have unique extensions of the same norm. Phelps’s theorem implies the Taylor-Foguel theorem (Herrero 2003, p. 88, Holmes 1975, p. 175) and was generalized by Park 1993.

The 2-ball sequence property mentioned in Sec. 7.4 provides a purely internal characterization of those $X$ whose every Hahn-Banach extension is unique—or, equivalently, of those $X$ with strictly convex dual—namely those $X$ in which every closed subspace satisfies the 2-ball sequence property.

8. Non-Archimedean Functional Analysis

By considering normed spaces $X$ over a field $F$ with an absolute value other than $\mathbb{R}$ or $\mathbb{C}$ we can glimpse what functional analysis looks like without the Hahn-Banach theorem. There is special interest in the case when the norm and absolute value are non-Archimedean, i.e.,

$$\|x + y\| \leq \max(\|x\|, \|y\|) \quad \text{for all } x, y \in X$$

Even in this context, a linear functional $f : X \to F$ is continuous if and only if it is bounded on the unit ball. Non-Archimedean analysis is quite similar to ordinary analysis in situations in which the Hahn-Banach theorem holds, quite different otherwise. Because mutually intersecting balls are concentric in non-Archimedean normed spaces, the binary intersection property simplifies to:

**Spherical Completeness.** Every decreasing sequence of closed balls has nonempty intersection.

It is similar in appearance to completeness—every decreasing sequence of closed balls whose diameters shrink to 0 has nonempty intersection—but stronger. $\mathbb{R}$ is spherically complete. Ingleton 1952 [cf. Narici et al. 1977] adapted Nachbin’s and Goodner’s arguments about the equivalence of extendibility and the binary intersection property to prove that a non-Archimedean Banach space is extendible if and only if it is spherically complete.

Pérez-García 1992 gives a thorough survey of the Hahn-Banach extension property in the non-Archimedean case, a situation in which Hahn-Banach extensions are never unique [Beckenstein and Narici 2004]. For the case when $F$ is not spherically complete, see Pérez-García and Schikhof 2003.

9. The Axiom of Choice

By teasing out a maximal element $F$ from the dominating extensions of $f$, the standard proof of the Hahn-Banach theorem (HB) uses the Axiom of Choice (AC) in the form of Zorn’s lemma.

9.1. Is HB $\iff$ AC? Does HB imply AC? as Tihonov’s theorem does? Can we call it “the analyst’s form of AC?” In a word: “No.” The details are as follows.

As is well known

$$\text{AC} \implies \text{Ultrafilter theorem (UT)}$$

namely that every filter of sets is contained in an ultrafilter. Halpern 1964 proved that UT $\implies$ AC. Łoś and Ryll-Nardzewski 1951 and Luxemburg [1962, 1967a,b] proved that UT $\implies$ HB. Pincus [1972, 1974] proved that HB $\nleftrightarrow$ UT. We therefore have the following irreversible hierarchy:

$$\text{AC} \implies \text{UT} \implies \text{HB}$$
The “prime ideal theorem for Boolean algebras” asserts that there is a function $F$ defined on the class of all Boolean algebras $B$ such that $F(B)$ is a prime ideal of $B$ for each $B$. Using techniques from non-standard analysis, Luxemburg 1962 showed that the prime ideal theorem implies the Hahn-Banach theorem and conjectured that the prime ideal and Hahn-Banach theorems might be equivalent. Halpern 1964, however, proved that the prime ideal theorem is strictly weaker than AC. Luxemburg 1967b showed that a modified form of the Hahn-Banach theorem is valid if and only if every Boolean algebra admits a nontrivial measure. The modification consists of allowing the extended linear functional on the real Banach space $X$ to take values in a “reduced power of the reals” (as used in nonstandard analysis) rather than $\mathbb{R}$; the modified version is also equivalent to the unit ball of the dual of the normed space $X$ being convex-compact in the weak-* topology, i.e., that every family of weak-* closed convex sets with the finite intersection property has nonempty intersection. Luxemburg and Váth 2001 proved that the assertion that any Banach space has at least one nontrivial bounded linear functional implies the Hahn-Banach theorem.

### 9.2. Avoiding AC

Various people have proved weaker versions of the theorem that do not rely on the Axiom of Choice. These include:

- Garnir, de Wilde and Schmets 1968 use only the Axiom of Dependent Choices—a little stronger than the countable axiom of choice but weaker than AC—to prove a Hahn-Banach theorem for separable spaces.
- Ishihara 1989 proved another ‘constructive’ version.
- Mulvey and Pelletier 1991. Locales generalize the lattice of open sets of a space without reference to the points of the space. Mulvey and Pelletier avoid dependence on AC. They use locales to prove a version of the Hahn-Banach theorem in any Grothendieck topos.
- Dodu and Morillon 1999 add a little and take a little. They suppose that the Banach space $X$ satisfies the stronger completeness requirement that Cauchy nets converge. They then prove the Hahn-Banach theorem for uniformly convex Banach spaces whose norm is Gateaux differentiable without AC. Still assuming that the Banach space $X$ satisfies the stronger completeness requirement, Albius and Morillon 2001 show that to have the Hahn-Banach theorem, it suffices to have a strengthened differentiability condition, uniform smoothness, namely, the uniform convergence of $(\|x + h\| + \|x - h\| - 2\|x\|)/\|h\|$ as $h \to 0$ for all $x$ on the surface of the unit ball of $X$.

### 10. “Sandwich Theorems” and Another Approach

Mazur and Orlicz 1953 used the Hahn-Banach theorem to prove interpolation theorems such as:

**Theorem 8.** Let $p$ be a sublinear functional and $q$ a superlinear functional on the real vector space $X$ such that $q \leq p$. Then there exists a linear functional $f$ on $X$ such that $q \leq f \leq p$.

In a survey of results on the existence of linear functionals satisfying various conditions, Lassonde 1998 proves a blend of the Banach-Alaoglu and Hahn-Banach theorems on a real vector space to deduce results on the separation of convex functions by an affine function.
König and Rodé and others 1968–1982 reversed the direction of the usual proofs of the Hahn-Banach theorem. To describe them, let $X$ be a real vector space and let $X^\#$ denotes the class of all sublinear (positive homogeneous, subadditive) functionals on $X$. Order $X^\#$ pointwise by $p \leq q$ if and only if $p(x) \leq q(x)$ for all $x \in X$. It happens that a sublinear functional $p$ on $X$ is linear if and only if it is a minimal element of $(X^\#, \leq)$. Given a subspace $M \subset X$ and a linear functional $f : M \to \mathbb{R}$, $f \leq p$, $p \in X^\#$, consider the collection of all sublinear functionals $q$ such that $f \leq q$ on $M$ and $q \leq p$ on $X$ and choose a minimal $q$ from this class. Any such $q$ is a linear extension of $f$ to $X$. Instead of enlarging $f$, they squash $p$. The method is considered at length in Narici and Beckenstein 1985 [Sec. 8.4].

Approaching the Hahn-Banach theorem by means of minimizing sublinear functionals leads to a variety of “sandwich theorems” such as Theorem 9 and its generalization below.

**Theorem 9.** Let $p$ be a sublinear functional defined on the real vector space $X$, let $S \subset X$ be convex and let $f : S \to \mathbb{R}$ be concave. If $f \leq p$ on $S$ then there exists a linear functional $F$ on $X$ such that $f \leq F$ with $F \leq p$ on $S$.

**Theorem 10.** [König 1982, Th. 2.1; cf. Narici and Beckenstein 1985, Ex. 8.202(b)] Let $p$ be a sublinear functional defined on the real vector space $X$, let $S$ be any subset of $X$ and let $f : S \to \mathbb{R}$. If $f \leq p$ on $S$ and there exist $a, b > 0$ such that

$$\inf_{w \in S} [p(w - au - bv) - f(w) + af(u) + bf(v)] \leq 0$$

for all $u, v \in S$ then there exists a linear functional $F$ on $X$ such that $f \leq F$ and $F \leq p$ on $S$.

Two nice surveys of this material are König 1982 and Fuchssteiner and Lusky 1981. Neumann 1994 simplifies some of the proofs of these results, develops some new ones and has a good bibliography on the subject as does Buskes 1993.

Rodé 1978 proved a very general version of the Hahn-Banach theorem, one flexible enough to apply to contexts other than linear spaces. König 1987 simplified Rodé’s proof. Páles 1992 offers two geometric versions [Theorems 1 and 2] of Rodé’s theorem. His Theorem 2 implies Rodé’s theorem, thereby providing another proof of it.

## 11. Locally Convexity and Hahn-Banach Extensions

By saying that a topological vector space has the **Hahn-Banach extension property** (HBEP) we mean that any continuous linear functional on any linear subspace possesses a continuous extension to the whole space. Every locally convex Hausdorff space has HBEP. What about the converse?

In the absence of local convexity, a topological vector space $X$ need not have any nontrivial continuous linear functionals at all. For $0 < p < 1$, for example, the dual of (the non-locally convex space) $L_p[0, 1]$ is trivial [Day 1940; Kalton et al. 1984]. Although local convexity is not essential for the existence of nontrivial continuous linear functionals, it helps: A topological vector space $X$ has a nontrivial dual if and only if there is a proper convex neighborhood of $0$ [Köthe 1969, p. 192; Kalton et al. 1984, p. 17].

A. Shields had observed that, given a dual pair $(X, X')$, any topology between the weak ($\sigma(X, X')$) and the Mackey topologies ($\tau(X, X')$) has HBEP and asked if such topologies had to be locally convex. Gregory and Shapiro 1970 showed that if $\sigma(X, X') \neq \tau(X, X')$ there are non-locally convex topologies in between,
thereby providing a plethora of non-locally convex topologies with HBEP. Kakol 1992 gives an elementary construction for an abundance of vector topologies \( \tau \) on a fixed infinite-dimensional vector space \( X \) such that \((X, \tau)\) does not have the HBEP even though \( X' \) is rich enough to separate the points of \( X \).

Topological vector spaces can have rich duals and still not have HBEP. For \( 0 < p < 1 \), the non-locally convex spaces \( \ell_p \) and the Hardy spaces \( H_p \) do not have HBEP but have an abundance of continuous linear functionals such as the evaluation functionals at \( n \in \mathbb{N} \) for \( \ell_p \) or points \( t \) in the open unit disk for \( H_p \). Indeed, \( \ell_p' = \ell_\infty \) for any \( 0 < p < 1 \) [Kalton et al. 1984]. Let us say that a subspace \( M \) of a TVS \( X \) has the separation property if any \( x \notin M \) can be separated from \( M \) by a continuous linear functional. If it is possible to extend any \( f \in M' \) to an element of \( X' \), we say that \( M \) has the extension property. For individual subspaces there is no connection between the separation and extension properties. Duren et al. 1969 showed that there are closed subspaces \( M \) of \( H^p, 0 < p < 1 \), with the separation property which do not have the extension property and vice-versa. Nevertheless (ibid.), for an arbitrary TVS \( X \), every subspace has the separation property if and only if every subspace has the extension property.

Shapiro 1970 showed that an \( F \)-space (complete metrizable TVS) \( X \) with a basis has the HBEP if and only if it is locally convex. Kalton removed the “with a basis” hypothesis. Using the fact that an \( F \)-space has HBEP if and only if every closed subspace is weakly closed and developing some basic sequence techniques for \( F \)-spaces, Kalton [1974; Kalton et al. 1984 p. 71] showed that an \( F \)-space with HBEP must be locally convex. This is false without metrizability, however—Any vector space \( X \) of uncountable algebraic dimension with the strongest vector topology (1) is not locally convex or metrizable but (2) has the Hahn-Banach extension property [Shuchat 1972].

References


Mathematics Department, St. John’s University, Jamaica, NY 11439, USA
E-mail address: narici@gmail.com