Topics on analytic sets

by

R. Kaufman (Urbana, Ill.)

Abstract. A survey on some relations between analysis and descriptive set theory.

Introduction. The topics are presented in four essays and concern both set theory and one or another kind of analysis: real variables, functional analysis, harmonic analysis. Besides the proofs of the main results, examples and curiosities are presented in no logical sequence.

Various approaches to descriptive set theory are presented in [5, 14, 16, 19].

I. Some analytic sets in function space

1. L denotes the metric space whose elements are Lebesgue measurable subsets of (0, 1), with metric \( d(A, B) = \lambda(A \cup B) - \lambda(A \cap B) \). A sequence \((A_1, A_2, \ldots)\), i.e., an element of \( L^\infty \), is called an I-sequence (infinite intersection) if there is an infinite sequence \( M \subseteq \mathbb{N} \) such that

\[
\lambda\left( \bigcap_{m \in M} A_m \right) > 0.
\]

The class of I-sequences in \( L^\infty \) is denoted \( \mathcal{I} \).

**Theorem.** \( \mathcal{I} \) is an analytic set in \( L^\infty \), but not Borel.

We give the proof that \( \mathcal{I} \) is analytic, although it is entirely elementary. Let \( J \) be the set of all infinite sets of natural numbers, and \( h \) the function defined on \( J \times L^\infty \) by the formula

\[
h(M, A_1, A_2, \ldots) = \lambda\left( \bigcap_{m \in M} A_m \right).
\]

By the continuity properties of the Lebesgue measure, we see that \( h \) is upper semicontinuous (using the usual product topology in both \( J \) and \( L^\infty \)). Now \( \mathcal{I} \) is the projection on \( L^\infty \) of the subset of \( J \times L^\infty \) defined by \( h > 0 \), a set of type \( F_\sigma \). Thus \( \mathcal{I} \) is indeed an analytic set.

To prove the more difficult assertion, we define measurable sets \( E(m_1, \ldots, m_k) \) for each \( k \)-tuple of natural numbers by the formula

\[
E(m_1, \ldots, m_k) = \{ x : m_j x \notin (0, 1 - 2^{-j}) \text{ modulo } 1, \ 1 \leq j \leq k \}.
\]
To abbreviate we write $E(o)$ when $o = (m_1, \ldots, m_k)$, and $I(o) = k$, the length of $k$. We call the sequences strings, and we say that a sequence $\sigma_k$ with $I(\sigma_k) \rightarrow +\infty$ converges to an infinite sequence $(m_1, m_2, \ldots)$ provided the coordinates of $\epsilon_k$ converge to those of the infinite sequence.

**Lemma.** Suppose that $(\sigma_k)_{k=1}^{\infty}$ is a sequence of distinct strings, and $\lambda(\bigcap_{k=1}^{\infty} E(\sigma_k)) > 0$. Then the sequence contains a subsequence $(\sigma_{k_n})$ with $I(\sigma_{k_n}) \rightarrow +\infty$, convergent to an infinite sequence $(m_1, m_2, \ldots)$.

**Proof.** The strings $\sigma_k$ have length $\geq 1$, and first elements $m_k$. Hence the set $B$ defined by $0 \leq m_k x \leq 1/2 \pmod{1}$ for all $p$ has positive Lebesgue measure. If the numbers $m_k$ were unbounded, then by the Lebesgue density theorem $\lambda(B \cap E) \geq \lambda(E)/2$ for every measurable set $E$, i.e., $\lambda(B) = 0$. Therefore there is an infinite sequence $S_i$ such that $m_k = m_i$ for $p \in S_i$. Then $I(\sigma_p) \geq 2$ for all but one $p \in S_i$, whence $\sigma_k = m_2$ along an infinite subsequence $S_2 \subseteq S_i$, etc.

Let $\mathcal{S}$ be the set of all trees of strings of natural numbers ([14, p. 112]). We enumerate the strings by any method into a sequence $(\sigma_k)$, and map $\mathcal{S}$ into $\mathbb{N}^\infty$ by the formula (here $T$ is a tree)

$$A_T(T) = \begin{cases} \emptyset & \text{if } \sigma_k \notin T, \\ E(\sigma_k) & \text{if } \sigma_k \in T. \end{cases}$$

(The result of the mapping is then a sequence $(A_T(T), A_T(T), \ldots)$.) The mapping is obviously continuous, and the Lemma shows that if $(A_T(T))_{T}^{\infty}$ is an $I$-sequence, then $T$ contains an infinite branch (using the customary language applied to trees). If $\{m_1, (m_1, m_2), \ldots, (m_1, m_2, \ldots, m_k, \ldots), \ldots\}$ all belong to $T$ then $(A_T(T))_{T}^{\infty}$ is an $I$-sequence because

$$\lambda(E(m_1, \ldots, m_k)) \geq \prod_{j=1}^{\infty} (1-2^{-j}) > 0.$$ 

The trees containing an infinite branch do not form a Borel set in $\mathcal{S}$ ([14, p. 113]), whence $\mathcal{S}$ cannot be a Borel set in $\mathbb{N}^\infty$.

2. Let $Y$ be the Cantor set, represented by sequences $(y_1, y_2, \ldots)$ with $y_m = 0$ or 1. A probability measure $\mu$ on $Y$ will be called an $I$-measure if there is an infinite sequence $M \subseteq \mathbb{N}$ such that the set defined by $\{y_m = 1 \text{ for all } m \in M\}$ has positive $\mu$-measure. The set $\mathcal{S}_\mu$ of $I_\mu$-measures is (as before) an analytic set in the $\mathbb{N}^\infty$-topology, and we will see in a moment that it cannot be a Borel set.

To do this, we define a mapping on $\mathbb{N}^\infty$ to probability measures on $Y$. It will be convenient to use sequences $(f_0, f_1, \ldots)$ of measurable functions, equal to 0 or 1 a.e., as a representation of $\mathbb{N}^\infty$. We simply write $f$ for such a sequence. The measure $\mu(f)$ is defined by

$$f(y_1, y_2, \ldots) dy = \int \prod_{j=1}^{\infty} f_j(s_j, s_2, \ldots, s_k, \ldots) dx$$

for every polynomial $p$. It is clear that this formula is also valid for bounded Borel functions $p$ defined on $Y$. Moreover, it is evident that $\mu(f)$ is an $I_\mu$-measure if and only if $f = (f_1, f_2, \ldots)$ is determined by an $I$-sequence in $\mathbb{N}^\infty$. Thus $\mathcal{S}_\mu$ cannot be a Borel set.

3. Here are some speculations about the class $\mathcal{S}_\mu$. Let us write $A$ (locally) for the class of all closed sets of the type $\{y_m = 1, m \in M\}$ for infinite sets $M \subseteq \mathbb{N}$, and $A^+$ for the probability measures annihilating all $A$-sets. Finally, let $A^\perp$ be the class of all closed sets $F \vee E$ such that $\mu(F) = 0$ for every $\mu$ in $A^+$. The class $A^\perp$ is of class PCA in $2^\mathbb{N}$. If, for example, we knew that $A^\perp$ is not of type CA, that would imply that $\mathcal{S}_\mu$ is not Borel (and considerably more).

4. Finally, we give an application to recent work of Kechris and Lyons [15] concerning measures on the circle $T$ of length $2\pi$. (It is considered as an additive group.) A Rajchman set in $T$ is defined by an infinite sequence $M \subseteq \mathbb{N}$ and an open subset $V \neq \emptyset$ of $T$ as follows:

$$\{x \in H(M, V) \mid m \notin V \forall m \in M\}.$$ 

It is proved in [15] that $H^1$—the probability measures annihilating all of the sets $H(M, V)$—is of class CA but not Borel. We obtain this result from the one above on the class $\mathcal{S}_\mu$ and some elementary considerations of Fourier analysis; we do not, however, write out all the details.

To each sequence $P \subseteq \mathbb{N}$ we attach a probability measure $\mu = \mu_P$ on $T$ by the formula

$$\mu(n) = \int e^{-\mu n} dm = \prod_{n \in P} (1 + \exp(-2\pi \mu n)) / 2.$$ 

Thus, when $P$ is empty, $\mu$ is the Lebesgue measure $(dt/2\pi)$ on $T$; in general, the support of $\mu_P$ is

$$E(P) = \left\{ \prod_{n \in P} \sum_{m=0}^{\infty} e_m 2^{-m} : e_m = 0 \text{ or 1, } e_n = 0 \text{ for } n \in P \right\}.$$ 

When $P$ is infinite, then $E(P)$ is a Rajchman set, because, when $k \geq 1$,

$$k \in P, \; t \in E(P) \Rightarrow 0 \leq 2^k t < \pi \pmod{2\pi}.$$

Our next objective is to explain why $\mu_P(H(M, V)) = 0$ unless $M$ and $P$ are related in a certain way. For this we need information about $\mu$. It is possible to pursue this in great precision [17] but what we need can be derived from the identities

$$\frac{1}{2} \{1 + \exp(-iu)\} |\cos(u/2)|, \quad -\infty < u < +\infty, \quad \prod_{m=1}^{\infty} \cos^2 m u = (\sin u)^2 / u.$$ 

Thus, if $r$ and $s$ are positive integers, and $2^r \geq |u| > 2^s$, then $|\prod_{m=1}^{\infty} \cos^2 m u| \leq 2^r |u|^{-1}$.

This leads us to define $n = \pm 1, \pm 2, \pm 3, \ldots$

$$\delta(n, P) = \inf (|k \log 2 - \log |n|| : k \in P).$$
We can then assert that if $n_j \to +\infty$ and $|\mu_j| \geqslant \varepsilon > 0$, then $\lim \sup \delta(n_j, P) < c(\varepsilon) < +\infty$. From this we shall deduce that $\mu_j(H(M, V)) = 0$ unless $\lim \sup \delta(m_j, P) < +\infty$, where $M = \{m_1 < m_2 \ldots < m_j < \ldots\}$. We follow [17] closely.

Let $h$ be continuous on $T$, $h = 1$ off $V$, $0 \leqslant h \leqslant 1$, and such that $h$ has mean value $\eta < 1$. If there is a subsequence $(m_l) \subseteq M$ such that $\delta(m_l, P) \to +\infty$, then $h(m_l) - \eta$ converges weakly to 0 in the space $L^q(d\mu)$, as we see with approximation of $h$ by polynomials. Any set $B$ on which $h(m_l) \geqslant 1$ for each $l$ must therefore satisfy $\mu(B) = 0$. In particular, we can conclude that $H(M, V)$ has $\mu$-measure 0.

We return to the Cantor set $Y$, but we represent it by sets $P$ of natural numbers. We write $P^*$ for the set obtained by squaring the elements of $P$ and $\mu_P$ for the measure attached to the set $P^*$ by the method introduced above.

To each probability measure $\sigma$ in $Y$ we associate the $w^*$-integral

$$r(\sigma) = \int \mu_P \sigma(dP),$$

a probability measure in $T$. We assert now that $r(\sigma)$ belongs to $H^S$ if and only if $\sigma$ is not an $L^q$-measure. (This proves at once that $H^S$ is not a Borel set.) To begin the proof of the assertion, we suppose that $\sigma$ is an $L^q$-measure. Then there is an infinite set $Q$ of natural numbers such that $\sigma(P^* : P \subseteq P) > 0$. For any $Q \subseteq P$, then $\mu(P^* \setminus Q^*)$ is concentrated in the Rajchman set $E(Q^*)$, whence (by integration) $(r(e))(E(Q^*)) > 0$, i.e., $r(e) \notin H^S$.

Conversely, we suppose $r(e)(H(M, V)) > 0$. Then for a certain Rajchman set $H(M, V)$. Writing $M = \{m_1 < m_2 \ldots < m_j < \ldots\}$ we see that $\lim \sup \delta(m_j, P^* \setminus Q^*) < +\infty$ for a set of $P^*$ in $Y$ of positive $\sigma$-measure. Since the elements of $P^*$ are perfect squares, and $(n+1)^2 - n^2 > 2n$, we see that the element $q$ of $P$ such that $|m_j| \log 2 - q^2 < c(P)$ is unique for large $j$. After vigorous applications of Ergovic’s theorem, we obtain an infinite set $Q$ such that $\sigma(P^* : P \subseteq P) > 0$. Then $\sigma$ is an $L^q$-measure. We have thus completed the proof that $H^S$ is not a Borel set.

II. Representations of PCA sets. This essay is based largely on the elegant work of Becker [1]; it is intended to illustrate an application of properties of $\mathbb{M}_0$-sets to a problem of set theory instead of the more usual direction from set theory to $\mathbb{M}_0$-sets. The special properties of this class are discussed below.

Let $(f_j)$ be a sequence in $C[0, 1]$, and $L(f_j)$ the set of functions $g$ in $C[0, 1]$ which are pointwise limits of subsequences of $(f_j)$. Then $L(f_j)$ is of type PCA in $C[0, 1]$ since a projection of a coanalytic set. (In [1] this class is called $\Sigma^1_1$.) To explain this, let $l$ be the set of strictly increasing sequences of natural numbers $\delta = \{n_1 < n_2 < \ldots < n_l < \ldots\}$ and $B$ the following subset of $J \times C[0, 1]$:

$$\{(\delta, g) : \lim_{l} f_n(x) = g(x) \text{ for all } x \in [0, 1]\}.$$

Then $B$ is coanalytic (its complement is analytic) and $L(f_j)$ is its projection in $C[0, 1]$. The striking converse of this is proved in [1]:

Every set of type PCA in $C[0, 1]$ can be realized as $L(f_j)$ for a certain sequence $(f_j)$.

Examination of the technique in [1] shows that the convergence is generally unbounded; the notion of bounded convergence is more tractable, since it coincides with weak convergence in the Banach space $C[0, 1]$. I learned of this problem from Alexander Kochris ([1], p. 164).

THEOREM. Let $\mathcal{P}$ be a set of type PCA in $C[0, 1]$. Then there is a sequence $(f_j)$ in $C[0, 1]$ such that $\mathcal{P} = \{f_j\}$.

(i) $\mathcal{P} = L(f_j)$.

(ii) Every convergent subsequence $(f_{n_k})$ is uniformly bounded.

I. A key role is played by the notion of $M_0$-set; a brief explanation and lemmas follow. The theory is treated much more thoroughly, but from a similar viewpoint, in Chapters I, II of [14].

Let $M_0$ be the class of measures $\mu$ on the torus $T$ such that

$$\mu(|n|) = \int |e^{-i\omega}| d\mu \to 0 \quad \text{as } |n| \to \infty.$$

(This class is often called the Rajchman class.) When $\mu \in M_0$, all measures absolutely continuous with respect to $\mu$ share the same property. Consequently it is usual to work with probability measures in $M_0$; in this chapter we denote this class of measures by $P_0$.

A Borel set supporting an element of $P_0$ is called an $M_0$-set; in this chapter and the next only closed $M_0$-sets are of interest.

From the sequence $\cos t, \cos 2t, \ldots, \cos nt, \ldots$ no subsequence can converge on a set of positive measure for any element $\mu$ of $P_0$. The proof of this follows from the remarks above: let $h$ be a bounded Borel function. Then one sees easily that $\int(h \circ \cos nt) d\mu \to 0$, $\int(h \circ \cos nt) d\mu \to 0$ if a subsequence $\cos_{n_k} t$ were to converge on a set of positive $\mu$-measure, its limit would be 0 a.e. $d\mu$ on that set; at the same time $\lim \int \cos^2 n_k t = 1/2$ a.e. This proves in passing that the measure $\mu$ has no jumps, i.e., is a continuous measure.

Let $\mu \in P_0$, let $\mu$ have closed support $E$, and let $\mu \geqslant 0$. Then as explained in [14], p. 294, $E$ carries a measure $\mu \in P_0$ such that $\mu \leqslant \mu$ (and $\mu$ and $\nu$ are mutually singular) and $|\mu - \nu| < \delta$. (Theorems of this type appear to originate with Iwasawa-Musatov.) By Luzin’s theorem, $\nu$ can be adjusted so that its closed support has $\mu$-measure 0 (this is a bit stronger than the relation $\mu \leqslant \nu$).

We can now obtain a stronger version of this, which will find repeated application.

LEMMA 1. Let $\mu$ belong to $P_0$, let $\mu$ have closed support $E$, and let $\delta > 0$. Then we can find measures $\lambda_j \in P_0 \quad (j \geqslant 1)$ with supports $E_j \subseteq E$ such that $|\lambda_j - \mu| < \delta$ and the sets $E_j$ admit open neighborhoods $V_j$ which are pairwise disjoint.

Proof. First we find supports $\mu \subseteq P_0$ with supports $E_j = B_{y_j} \subseteq \ldots \subseteq B_{y_j} \subseteq \ldots$ such that $|\lambda_j - \mu| < \delta/3, |\lambda_j - \mu| < \delta/6, \ldots$, and $B_{y_j}$ has $\mu$-measure 0. (This is possible by the remarks above.) Then $|\lambda_j - \mu| < 2\delta/3$ and $\nu_j$ is concentrated in the set $B_{y_j} - B_{y_{j+1}}$. Hence we can find a measure $\lambda_j \in P_0, |\lambda_j - \mu| < \delta$, and the closed support $E_{j+1}$ of $\lambda_j$ is a subset of $B_{y_j} - B_{y_{j+1}}$. Thus $E_{j+1}$ is disjoint from the closure of the remainder set $\bigcup_{k=j+1}^{\infty} V_k$, so the open sets $V_j$ exist.
Lemma 2. We can find a system of open sets \( V(m_1, \ldots, m_k) \), defined for all strings of natural numbers \((m_1, \ldots, m_k)\) of lengths \( k = 1, 2, 3, \ldots \), such that

i. \( V(m_1) \supseteq V(m_1, m_2) \supseteq V(m_1, m_2, m_3) \supseteq \ldots \)

ii. \( V(m_1, \ldots, m_k) \cap V(m_1, \ldots, m_k, m_{k+1}) = \emptyset \) unless \( m_1 = m_2, \ldots, m_k = m_{k+1} \).

iii. Each intersection \( V(m_1) \cap V(m_1, m_2) \cap V(m_1, m_2, m_3) \cap \ldots \) contains a (compact) \( M \)-net of \( (m_1, m_2, \ldots, m_k) \).

Proof. We begin with any element \( \mu \in P_0 \), with support \( E \). We find disjoint open sets \( V(m_1) \), closed subsets \( V(m_1) \), and measures \( \mu(m_1) \) in \( E(m_1) \) such that \( |\mu(m_1)| < 1/2 \). (Clearly we want \( \mu(m_1) \in P_0 \).) For each \( E(m_1) \), we find disjoint open sets \( V(m_1, m_2) \), closed sets \( E(m_1, m_2) \), and so forth. Now we require that \( \mu(m_1, m_2) \leq 1/4 \), etc. We can also require that \( V(m_1, m_2) \subseteq V(m_1) \), since \( V(m_1, m_2) \subseteq V(m_1) \subseteq V(m_1) \). The last step is to observe that each sequence \( \mu(m_1), \mu(m_1, m_2), \mu(m_1, m_2, m_3), \ldots \) converges uniformly and the limit is the Fourier-Stieltjes transform of an element of \( P_0 \), whose support is a subset of \( E(m_1) \cap E(m_1, m_2) \cap \ldots \).

To agree with the machinery in [1], the open sets \( V(m_1) \), \( V(m_1, m_2) \), \ldots \) must have Lebesgue measures whose sum is finite. This is easy to arrange if we choose the initial \( M \)-net to have Lebesgue measure \( 0 \). After this we have to transform our sets from the torus \( T = [0, 2\pi] \) to the interval \( [0, 1] \). One presents no difficulty if we choose \( E \) in \([0, \pi]\).

2. The exposition in [1] leaves no room for improvement, so we simply note a few changes necessary to obtain (ii) in the Theorem. The definition of \( \mathcal{O} \) [1, line 2, p. 167] is expanded to include sequences such that \( |P_n| \supseteq 1 \). In 3.3(b), p. 165, instead of \( f = n \) we define \( f(s) = \cos 2\pi ns \) on the closed set \( \mathbb{P} \).

A natural question concerns the special position of the \( M \)-nets. Can this class be replaced by the class \( L^0 \) of sets of positive Lebesgue measure, or the class \( \Pi \) of perfect sets? Here \( L^0 \) is excluded because any disjoint subclass can be measurable. As for \( \Pi \), any bounded sequence \( (f_n) \) in \( C[0, 1] \) contains a subsequence convergent everywhere on some perfect set. Since this work was first completed, substitutes for the \( M \)-nets have been found, but all rely on some variant of the Riemann-Lebesgue Lemma.

3. The Banach space \( l^2 \) is an inseparable dual, and the weakly convergent sequences converge in norm—the so-called Schur property. Hence, each set of type \( L(f) \), the set of weak limits of subsequences of a fixed sequence in \( l^2 \), is closed. It would be interesting to find other spaces \( X \) in which some of the sets \( L(f) \) are not analytic even though \( X \) is less complicated than \( C[0, 1] \). They cannot be excluded that sets \( L(f) \) in \( X \) are always analytic but sometimes not Borel. Some applications of descriptive set theory to Banach spaces are presented by Szlenk [20] and Bourgain [3].

III. On functionals attaining their norm. Let \( X \) be a real Banach space and \( C \) a closed, bounded, convex set in \( X \). The continuous linear functionals on \( X \) (i.e. elements of \( X^* \)) which attain their suprema on \( C \) are called support functionals of \( C \) and form a set \( \mathcal{S}(C) \). (It would be more accurate to include the space \( X \) in the notation. Concerning \( \mathcal{S}(C) \) there are two classical theorems.

(a) James' Theorem [11, 12, 13, 7, 8]. \( \mathcal{S}(C) \neq X^* \) unless \( C \) is weakly compact.

(b) Theorem of Bishop-Phillips [2, 7, 4]. \( \mathcal{S}(C) \) is a norm-dense subset of \( X^* \).

There is a third case, equally natural: when \( \mathcal{S}(C) \) of the second category (i.e. not of the first category) for every c.b.c.s. \( C \). This occurs precisely when \( X \) has the RNP (Radon-Nikodým property). This belongs to a theory created in large part by Phelps, Stegall, and Bourgain [4].

When \( C \) is the unit ball of \( X \), the set \( \mathcal{S}(C) \) will be written \( \mathcal{S} \) – the class of linear functionals "attaining their norm". In the case of separable spaces, \( \mathcal{S} \) is an analytic set in \( X^* \), provided with the weak-topology.

We give a proof of the last assertion, because the method of proof has an interesting consequence.

Let \( A \) be the following subset of \( X \times X^* \): \( A \) contains pairs \((x, f)\) such that \( x \in X, \|x\| \leq 1, f \in X^*, \|f\| \leq 1, f(x) = \|f\| \).

When \( X \) is equipped with the norm topology and \( X^* \) with the weak-topology, this set is closed. When \( X \) is separable, the unit ball of \( X^* \) is compact and metrizable. The projection of \( A \) into \( X^* \) is just the part of \( \mathcal{S} \) in the unit ball of \( X^* \). Thus \( \mathcal{S} \) itself is analytic. (All of this is elementary.) If \( f \neq 0 \) and \( X \) is strictly convex then the equality \( f(x) = \|f\| \) has at most one solution \( x \) in the unit ball of \( X \). From a theorem of Lusternik [16, p. 397], it follows that \( \mathcal{S} \) is a Borel set. (\( X \) is strictly convex if the inequality \( \|x_1 + x_2\| < \|x_1\| + \|x_2\| \) is strict whenever \( x_1 \neq 0 \) and \( x_2 \) are linearly dependent. When \( X \) is not strictly convex, then some linear function \( f \neq 0 \) attains its norm on a line segment in the unit ball of \( X \).)

Every separable space \( X \) can be renormed to become strictly convex. The set \( \mathcal{S} \) for the new norm is then a Borel set; but we do not know if it can always be made into a \( G_\delta \)-set (i.e. for every separable space). Some spaces have no equivalent strictly convex norm (Day, cf. [7]). For the classical spaces \( l^p, c_0, C[0, 1], \ldots \), \( \mathcal{S} \) is a Borel set of rather low complexity. Nevertheless:

Theorem. \( \mathcal{S} \). Let \( X \) be a Banach space, separable but not reflexive. Then there is a norm \( \|x\| \) on \( X \) such that \( \|x\| \leq \|x\|_2 \leq 2 \|x\| \), and relative to the new norm \( \mathcal{S} \) is not a Borel set in \( X^* \).

More exactly, there is a norm-compact set \( K \) in \( X^* \) such that \( K \cap \mathcal{S} \) is not a Borel set in \( K \).

(The remark about the set \( K \) is provoked by a controversy in set theory related to Martin's axiom [18]. The more precise statement is true and has a definite meaning even when \( X \) is not separable, but we do not dwell on this point here.)

Because of the intricacy of the most general form, we present elementary examples for \( X = C[0, 1] \) and \( X = c_0 \). The first example illuminates some classical ideas in real analysis; the second does the same for Fourier analysis; and the most general relies on a device from stochastic processes.

First example, \( X = C[0, 1] \). Let \( X_1 \) be the subset of \( X \) defined by the conditions \( f(0) = 0, \ Var(f) \leq 1 \).
Here, \("Var"\) means total variation. Later, we use \(B_r(X)\) to mean the ball \(|x| \leq r\) in \(X\). Now \(Y_2\) is a closed, bounded, convex subset of \(X\). The space \(X^*\) has unit ball equal to 
\[ \mathcal{C}(B_r(X) \cup Y_2). \]
Fortunately, the argument does not depend on an exact description of this set, only an approximation. Before proceeding to an explicit description of the dual norm in \(X^*\), we write two elementary but important facts about Stieltjes integrals.

(A) When \(f \in Y_1\), then the Stieltjes measure \(df\) is continuous, i.e. has no jumps. In fact, since \(f\) has bounded variation, \(f = f_1 - f_2\), where \(f_1\) and \(f_2\) are monotone increasing. Now \(f_1\) and \(f_2\) have exactly the same jumps (left-hand jumps or right-hand), and these can be canceled so that \(f = g_1 - g_2\), where \(g_1\) and \(g_2\) are monotone and continuous. Then \(df = dg_1 - dg_2\).

(B) For every \(f\) in \(X\) and every measure \(\mu\) in \(I = [0,1]\)
\[ \int f \mu = \int (f(1) + f(0)) t \mu(\{0,1\}) dt \]
\[ f \in Y_2 \Rightarrow \int t \mu(\{0,1\}) dt = \int f \mu. \]
To see this we can write \(f(t) = c(\{t,1\})\) with a continuous Borel measure \(c\) in \((0,1)\), and then apply Fubini's theorem.

To write the formula for the dual norm \(||\mu||^\ast\) we define first
\[ ||\mu||^\ast = \sup \{ ||(0,1)\| \mu(\{0,1\}) < t < 1 \}. \]
Then we have

**Lemma 1.** For every measure \(\mu\), \(||\mu||^\ast = \max\{1, 2||\mu||\}.\)

**Proof.** From the formula in (B) we get
\[ \int f \mu \leq 2 ||\mu|| \quad \text{for any } f \in Y_1. \]
The definition of the unit ball in \(X^*\) shows that the right member in the lemma is no less than \(||\mu||^\ast\) and, of course \(||\mu||^\ast < ||\mu||\).

It remains to be verified that \(||\mu||^\ast > 2||\mu||\). Suppose then that \(||\mu||^\ast > c > 0\); since \(\mu(\{0,1\})\) is continuous on the right, there is an open interval \(U \subset (0,1)\) on which \(\mu(\{0,1\}) > c\). Let \(\chi\) be the characteristic function of \(U\) and let \(g\) be the absolutely continuous function such that \(g(1) = 0\) and \(g'(t) = \mu(\{0,1\}) \chi(t)\) a.e. Then \(g\) has variation at most \(a = m(U)||\mu||\), whence \(2a^{-1} g\) belongs to \(Y_1\). Moreover, \[ \int g \mu = -\mu(\{0,1\}) g \mu) \quad \text{and} \quad \int g d\mu \leq c^2 m(U). \]
Thus
\[ ||\mu||^\ast \geq c^2 m(U) a^{-1} = 2e^{c/2} ||\mu||. \]
As this is true for any \(c < ||\mu||\), we obtain \(||\mu||^\ast \geq 2||\mu||\), and Lemma 1 is proved.

**Lemma 2.** Suppose that \(||\mu||^\ast = 1, ||\mu|| < 2\). Then \(\mu\) attains its norm on the unit ball of \(X^*\) if and only if the set \(\{\mu(x) = ||\mu||\}\) is uncountable.

**Remark.** Lemma 2 depends on the correctness of remarks (A) and (B), in contrast with the habitual latitude in real analysis.

**Proof.** Suppose that \(||\mu||^\ast = 2\), and \(f\) belongs to the convex hull of \(B_1(X) \cup Y_2\).
Then \(f = \lim (t_0, g_0 + (1-t_0) h_0)\) with \(0 < t_0 < 1\), \(g_0 \in B_1, h_0 \in Y_2\). Then \(f = \lim \mu(\{0,1\}) + (1-t_0) \mu(1-2 + t_0||\mu|| - 2) \Rightarrow \mu(\{0,1\}) \leq 1\) on \((0,1)\). Consequently, \(t_0 \to 0\) and \(f = \lim h_0\); and \(Y_2\) is closed, whence \(f \in Y_2\). (This deduction is used several times later.) Thus
\[ -\mu(\{0,1\}) df(t) = 2 + \int (f^0) df(t) \leq 2, \quad \mu(\{0,1\}) < 1 \quad \text{on } (0,1). \]
Combining the equality and the inequalities we see that the measure \(df\) is concentrated in the set where \(\mu(\{0,1\}) = \pm 1\), and the latter set must be uncountable.

Conversely, if \(A(\mu)\) is uncountable, it contains a perfect set \(A_0\) because it differs from its closure by at most a countable set. Let \(\lambda\) be a continuous probability measure in \(A_0\) and let \(f\) be defined by the equations \(f(t) = 0, df = \mu(\{0,1\}) dt\). Then \(Var(f) < 1\), so that \(f\) belongs to \(Y_2\) and \(\int df = 2\). That is, \(\mu\) attains its norm 2 on the unit ball of \(X^*\).

To finish the example, we use a classical theorem of Hurewicz [10; 14, p. 137]: in the set \(F\) of non-voyed closed subsets of \([3/4,1]\) on the uncountable closed sets form an analytic, non-Borel set. We now define for each element \(S\) of \(F\) a function \(\varphi(t) = \varphi(t, S)\) in \(C^1[0,1]\) such that
\[ \begin{align*}
& (a) \quad 0 < \varphi(0) < 1, \\
& (b) \quad \varphi(0) = 0, \varphi = 1 \text{ on } [0,1], \varphi < 1 \text{ on } [0,1] - S, \\
& (c) \quad \int \varphi' dt < 3/2, \\
& (d) \quad \text{the map from } S \text{ to } \varphi(t, S) \text{ is a continuous map into } C^1[0,1].
\end{align*} \]
The measure \(\mu\) defined by \(\mu(\{0,1\}) = \varphi(t, S)\) attains its norm precisely when \(S\) is uncountable, and all the elements \(\mu\) belong to a normal-compact set \(X\) in the dual space. Thus \(K \cap \mu\) is not a Bor.

The following construction of \(\varphi(t, S)\) is perhaps excessively detailed; many variants would serve equally well. Let \((I_n)\) be a sequence of intervals in \((-1,2)\), forming a basis for the open subsets of \((-1,2)\). Let \(u_n \in C^1(-\infty, \infty), u_n > 0\) on \(I_n\), \(u_n = 0\) outside \(I_n\), and \(\|u_n\| + \|u_n'\| < 2^{-n/3}\). We define first
\[ \psi(t, S) = 1 - \sum_{n=1}^\infty d(S, I_n) u_n(t), \quad 0 \leq t \leq 1. \]
This fulfills all of (a)–(d) except \(\varphi(0) = 0\). We define (at last)
\[ \varphi(t, S) = \begin{cases} \\
\psi(t, S), & 0 < t \leq 1, \\
\sin \psi(t, S), & 0 \leq t \leq 1/2.
\end{cases} \]
Second example, \(X = C_0\). To avoid certain technicalities, we use complex scalars. The space \(C_0\) is realized as trigonometric series \(\sum_n \omega_n e_n(0)\), where \(e_n(0) = e^{i n}\), and the dual space as absolutely convergent trigonometric series \(\sum_n e_n\). The duality
is defined by $(e_0, e_0) = 1$ if $m+n = 0$, $(e_n, e_n) = 0$ otherwise. (The advantage of this formalism will appear presently.) The set $Y$ consists of those elements of $X$ which are Fourier-Stieltjes series of measures of norm at most 2. These are the series $\sum a_i e_i$ such that $\| a_i b_i \| \leq 2 \sup |\lambda| e_{\lambda}(i)$ for all series $\sum b_i e_{\lambda}(i)$ in the dual. We observe that when $\sum a_i e_i$ represents a measure $\mu$ and $f$ belongs to the dual, $\langle \sum a_i e_i, f \rangle = \int d\mu$, a Lebesgue integral. As before the norm $\| f \|$ has for its ball the closed convex hull of $B_i \cup Y$, and the dual norm is written $\| f \|_\bullet$.

**Lemma 1.** Let $f(t) = \sum a_i e_{\lambda}(t)$ be an absolutely convergent Fourier series. Then $\| f \|_\bullet = \max \{ \| f \|, 2 \| f \|_1 \}$.

**Proof.** It is clear that $\| f \|_\bullet \geq \| f \|$. Suppose that $\| f \| > c$, so that $\| f \| > c$ on some open interval $I$ on the circle, of length $m(t)$. Let $\mu$ be the measure $\int (t) dt$, so that $\mu$ has variation at most $\| f \|_1 m(t)$. By the Riemann–Lebesgue Lemma, the Fourier-Stieltjes series of $\mu$ belongs to $X$, and (identifying $\mu$ with that series) $\mu \in \mathcal{M}$, with $a = \| f \|_1 m(t)$. Now $(\mu, f) \geq m(t) \| f \|_1$, whence $\| f \| \geq 2c \| f \|_1 - 2c$. Thus we get $\| f \| \geq 2c$; the inequality $\| f \|_\bullet \leq \| f \|_1, 2 \| f \|_1$ is obvious from the definition of $Y$.

**Lemma 2.** Let $f$ belong to the dual space, and $\| f \|_1 = 1$. Then $f$ attains its norm, $\| f \|$, on the unit ball of $X^*$ if and only if the set $A(f)$ of $t$ such that $|f(t)| = 1$ is an $\mathcal{M}_d$-set. ($\mathcal{M}_d$-sets were introduced at the beginning of Section II.)

**Proof.** If it attains its norm $= 2$, then $\langle a, f \rangle = 2$ for some series $\sum a_i e_{\lambda}(t)$ in $Y$, by an argument presented in the first example. Now the series represents a measure $\mu$ of variation at most 2, whence $\| \mu \| = 2$. Thus $\| f \| = 1$ a.e. $dt$, that is, $\mu$ is concentrated in $A(f)$, and $A(f)$ is a $\mathcal{M}_d$-set. Conversely, if $A(f)$ is a $\mathcal{M}_d$-set, it carries a probability measure such that $\int (\pm \infty) = 0$. The same property is shared by $\mu = 2f^* \lambda$, an element of $Y$. But then $\langle f, \lambda \rangle = 2$, since $f^* = 1$ a.e. $dt$.

We use the fact that the class of $\mathcal{M}_d$-sets is not a Borel set in the class of all closed sets in the circle [14]. We map each closed set $\mathcal{S}$ of $\theta$ to a smooth function $\varphi$ such that $0 \leq \varphi \leq 1$ and $A(\varphi) = \mathcal{N}$. A small change from the construction done before is that all functions must be periodic. By Parseval's formula, periodic functions with bounded derivatives have absolutely convergent Fourier expansions. In order to ensure that the norm of $f$ in the dual space is less than 2, we simply replace $f$ by $1 - e + e$, for a suitable constant $e > 0$. Since the embedding of $C^1$ into $C^\theta$ is compact (again by Parseval), we obtain a perfect analogue of the first example.

Recent work on descriptive set theory of classes occurring in harmonic analysis is presented by Tardivel [21] and Godefroy [9].

The general case of Theorem $\mathcal{A}$. Like the two special cases already presented, this depends on an analytic set that is not Borel, in an auxiliary metric space. Let $J$ be the Polish space of all infinite sets of natural numbers, and let $\mathcal{B}$ be a closed subset of $J$. Then $\mathcal{B}^+$ denotes the class of sets containing some member of the class $\mathcal{B}$. Thus $\mathcal{B}^+$ is always analytic in $J$.

**Lemma 1.** For a certain closed $\mathcal{B} \subset J$, $\mathcal{B}^+$ is not a Borel set.

**Proof.** An explicit example is the collection of all sequences

1, 2, ..., 2, 3, ..., 4, 5, ..., ...

where $n_1 \geq 1, n_2 \geq 2, n_3 \geq 1, \ldots$. The assertion that $\mathcal{B}^+$ is not Borel is very close to the same assertion about the set of trees containing an infinite branch [14, p. 113].

Before the next lemma we present a summary of some classical theorems of functional analysis. These are used to prove Lemma 2 below, and also Lemma 1 in IV.

(1) Suppose $X$ is a Banach space, $x_1, \ldots, x_n$ elements of $X^*$, $x \in X^*$, and $c > 0$. Goldstine's Theorem asserts that there is some $x$ in $X$ such that $\| x \| = \| x^* \|$ and $\| x_i^* - x^* \| < c$, $1 \leq i \leq n$. Instead, we suppose $\| x \| = r$. Then, by successive approximation, we find $x$ in $X$ such that $\| x \| < r$ and $\| x_i^* - x^* \| < \varepsilon$, $1 \leq i \leq n$.

(2) Let $x \in X^*$, and let $L$ be a finite-dimensional subspace of $X$. As usual $L^\perp$ is the null-space of $X$ in $L$. Then $\| x \| = \| x \|_L$ for each $x$ in $L$. Then (Hahn–Banach) there is some $g$ in $X^*$ such that $\| g \| = r$ and $g \neq f$ on $L$. By linear algebra, $\| f \| = \| f \|_L$.

**Lemma 2.** Let $X$ be a Banach space that is not reflexive. There are sequences $(x_i)_n$ in $X$ and $(\xi^*_i)_n$ in $X^*$ such that

(i) $\| x_i \| < 1$, $\| \xi^*_i \| < 1$,

(ii) $\| x_i \| = 1$ when $1 \leq i \leq k$, $\| x_i \| < 1$ when $1 \leq i < k$.

**Proof.** Let $\xi$ be $X^*$, $\| \xi^*_i \| < 4/3$, and $d(\xi^*_i, X) > 1$, and let $f_i \in X^*$, $\| f_i \| < 1$. $\xi(\xi^*_i) = 1$. Again, let $x_i \in X$, $\| x_i \| < 4/3$, $f_i(x_i) = 1$. Now, since $d(\xi^*_i, X) > 1$, there is some $x_i \in X$, $\| f_i \| < 4/3$, such that $f_i(x_i) = 0$ and $\xi(x_i) = 1$. (In the opposite case the distance from $\xi$ to $X$ would be at most 1) Since $\| \xi^*_i \| < 4/3$, there is some $x_i \in X$, $\| x_i \| < 4/3$, such that $f_i(x_i) = \xi(x_i) = 1$, and $f_i(x_i) = 1$.

**Remark.** A similar lemma occurs in James' Theorem. Any weak accumulation point $y$ of the sequence $(x_i)$ would satisfy $f_i(y) = \xi(y) = \xi(x_i) = 1$, and therefore $y$ could not belong to the linear span of $(x_i)$. This contradiction shows that some separable subspace of $X$ is not reflexive.

Returning now to the metric space $J$, we write its elements as sequences $\eta = (n_1, n_2, \ldots)$ so that the $r$th element $n_r$ depends continuously on $\eta$. We define a homeomorphism $\psi$ of $X$ into $J$ by the formula

$$\psi(\eta) = \sum_{r=1}^\infty \eta_r x_r, \quad \eta = (n_1 < n_2 < \ldots$$

We observe the implication

$$f_i(\psi(\eta)) < 2^{-r} = n_r < k.$$
To explain Lemma 3, we observe that $J$, being a complete, separable metric space, is a $G_δ$-set in a compact metric space. Thus the notion of a probability measure in $m$ has a definite meaning. Since $ψ$ is continuous on $m$, the integral can be taken as a strong (Bochner) integral.

Proof. Every element $x$ of $A$ is a limit $\lim_{n} \int (\psi (d) d\lambda_{x})$ of a sequence $(\lambda_{x})$ of discrete probability measures in $m$. For each $\delta > 0$ we proceed to find a compact set $K_{\delta} \subseteq m$ such that $\lambda_{x} (K_{\delta}) > 1 - \delta$ for every $m$. This is expressed by saying that the sequence of measures is "uniformly tight". We observe first that for each $m = 1, 2, 3, \ldots$

$$\lim_{n} \int (\psi (d) d\lambda_{x}) = 0,$$

whence there is some $k_{n}$ such that $\lambda_{x} (K_{\delta}) < 4^{-n} \delta$ for all $m$, and $s = 1, 2, 3, \ldots$. On the set of $x$ such that $n_{s} \geq k_{s}$, the inequality $\lambda_{x} (\psi (d)) \geq 2^{-s}$ holds. Since $\lambda_{x} (\psi (d)) > 0$ is always true, the set $\bigcap_{s=1}^{\infty} \{x: n_{s} \geq k_{s}\}$ has $\lambda_{m}$-measure $\geq 1 - \Sigma_{s=1}^{\infty} 2^{-s} = 1 - \delta$, for every $m$.

Thus, some subsequence of $(\lambda_{x})$ converges $\omega$ to a measure $\lambda$ concentrated in $m$.

Since $\psi$ is continuous on $m$, we obtain $x = \int (\psi (d) d\lambda)$.

Next we define a map $Φ$ of $J$ into $X^*$. To prevent confusion between $Φ$ and $ψ$, we write this map as $Φ (E)$. First, let

$$\varphi_{E} (1) = 1 \quad \text{for any set } E,$$

$$\varphi_{E} (m) = \varphi_{E} (m+1) \quad \text{if } m \in E, \ m \geq 2,$$

$$\varphi_{E} (m) = 1 - 5^{-m} \quad \text{if } m \notin E, \ m \geq 2.$$

Next we choose numbers $a_{E}$ so that the series $Φ (E) = \Sigma_{s=1}^{\infty} a_{E} (s) f_{s}$ satisfies

$$Φ (E) (x) = \varphi_{E} (m), \ m = 1, 2, 3, \ldots$$

Explicitly $\Sigma_{s=1}^{\infty} a_{E} = \varphi_{E} (m)$, or $a_{E} (1) = 1$, $a_{E} (2) = \varphi_{E} (2) - \varphi_{E} (1)$, $a_{E} (3) = \varphi_{E} (3) - \varphi_{E} (2)$, ... Thus, $a_{E} (1) = 1$, $|a_{E} (2)| \leq 5^{-1}$, $|a_{E} (3)| \leq 5^{-2}$, ..., whence $|Φ (E) (x)| \leq 5/4$, and the range of $Φ$ has compact closure in $X^*$.

Lemma 4. (a) For every $E$ and every $n$, $Φ (E) (ψ (d)) \geq 0$, and the supremum of $Φ (E)$ over $A$ is $+1$.

(b) The supremum is attained if and only if $E^* = E \cup \{1\}$.

Proof. The first statement of (a) is a consequence of the inequality $|φ| \leq Φ (ψ (d)) \leq 1$. For the second, we observe that $m$ contains sequences $(1, n_{1}, n_{2}, \ldots)$ whose second element $n_{1}$ is as large as we please. For every set $E$ we have $Φ (ψ (d)) > 1 - 5^{-n_{1}}$.

For (b) we observe that $Φ (E) (ψ (d)) = 0$ if and only if the set $n$ (actually, a sequence) is contained in $E \cup \{1\}$. The representation of $A$ established in Lemma 3 yields the assertion.

Since the functions $Φ (E)$ satisfy $|Φ (E)| \leq 5/4 < 4/3$, we conclude that $|Φ (E)| = 4/3$. By Lemma 4 we see that $Φ (E)$ attains its norm $+4/3$ if and only if it attains the value $+1$ on $A$, and only if $E^* \subseteq m$.

The next step is to confirm that the last condition defines a non-Borel subset of $J$.

Let $J_{1}$ be the subset of $J$ defined by $E = E^*$, i.e., $n_{1} = 1$. Since $J_{1}$ is closed, and $J_{1} \subseteq J_{2}$, it is clear that the set of $E$'s defined by $E^* \subseteq m$ is not a Borel set. Since the functionals $Φ (E)$ belong to some norm-compact set $K$ in $X^*$, $K \cap m$ is not a Borel set in $X^*$.

IV. A variant of Theorem $\mathcal{A}$. This section uses a bit more about weak and weak* topologies than III. It shows how the mappings $ψ$ and $Φ$, after small adjustments, can be combined with imperfect analogues of Lemma 2 of III. Suppose that $X$ is not reflexive, and therefore $X^*$ is not. Then some norm $|| \cdot ||$ in $X^*$ fulfills the precise version of Theorem $\mathcal{A}$ (which applies to non-separable spaces). But $|| \cdot ||$ is not always the dual of a norm in $X$; according to an interesting theorem of Davis and Johnson [S], $X^*$ can be renormed so that it is not isometric with any dual space! (A key idea is due to Kadec.) Nevertheless:

**Theorem $\mathcal{A}$.** The space $X$ carries a norm $|| \cdot ||^*$ such that $||x||^* \leq ||x|| \leq 2 ||x||^*$, and the dual norm in $X^*$ fulfills the precise version of Theorem $\mathcal{A}$.

(Of course $||x||^* \leq ||x|| \leq 2 ||x||^*$ for all linear functionals $x^*$.)

Before the first lemma, we make some remarks about the isometric imbeddings $Z \subseteq Z^{**}$ and $Z^* \subseteq Z^{***}$. Since the first of these is a proper inclusion, there is a functional $ξ : Z^{**} \rightarrow Z$ such that $||ξ|| = 1$ and $ξ = 0$ on $Z$. We claim that $d (ξ, Z^*) \geq 1/2$. Otherwise $||ξ - f|| < 1/2$ for some $f \in Z^*$. In particular, $||ξ|| \leq 1 ||x||$ for every $x \in Z$, since $ξ (x) = 0$. But then $||ξ|| \leq 1/2$, whence $||ξ - f|| \geq 1/2$.

The inequality $d (ξ, Z^*) \geq 1/2$ means that for each finite-dimensional subspace $F$ of $Z^*$, and each $ξ \geq 0$, there is some $y$ in $Z^{**}$ such that $y = 0$ on $F$, $||y|| = 1$, $ξ (y) \geq 1/2 - 

**Lemma 1.** Let $Z$ be separable, but not reflexive. Then there are sequences $(f_j) \subseteq Z^{**}$, $(y_j) \subseteq Z^*$ such that

(i) $||y_j|| \leq 5/4$, $||f_j|| \leq 2$ for $k \geq 2$,

(ii) $f_j (y_j) = 1$ when $1 \leq k \leq f_j$, $f_j (y_j) = 0$ when $1 \leq j < k$,

(iii) $\varphi_{E} (m) = \varphi_{E} (m+1) \quad \text{if } m \notin E, \ m \geq 2.$

Proof. Let $(z_j)$ be a dense sequence in $Z$, and let $ξ \in Z^{***}$, $||ξ|| = 5/4$, $ξ = 0$ on $Z$ (considered as a proper subspace of $Z^{**}$). Then there is some $f_j \in Z^{**}$ such that $||f_j|| \leq 1$ and $ξ (z_j) = 1$. Thus, we can choose $y_j \in Z^*$ such that $||y_j|| \leq 4/3$ and $f_j (y_j) = ξ (z_j) = 1$. Next, we take $(z_j) \subseteq Z^{**}$ such that $||z_j|| \leq 5/6$, $z_j (ξ) = 0$. Then, there is an $f_j \in Z^{**}$ such that $||f_j|| < z_j (ξ) = 0$, $ξ (f_j) = 1$. In the opposite case, the distance from $ξ$ to $y_j$ would be at most $1/2$. Finally, there is a $y_j \in Z^*$ such that $||y_j|| \leq 4/3$, $y_j (ξ) = ξ (f_j) = 1$, $y_j (f_j) = ξ (ξ) = 1$, etc. (At the third step we choose $f_j$ so that $||f_j|| < 2$, $f_j (y_j) = ξ (f_j) = 0$. The $f_j$ is $1$.)

...
Next, we need a small change in the definition of $\varphi_k$ and $\Phi(E)$. Let $\varphi_k$ be defined so that

$$\varphi_k(m) = 1 - 2^{-m-2} \quad \text{when } m \geq 2, \ m \notin E.$$  

Then the $\Phi(E)$ constructed by the same method as before satisfies $|\Phi(E)| < 5/4$.

We can now complete the proof of Theorem $\alpha$ in the special case of separable spaces $X$, i.e. taking $Z = X$. Here we choose for $A^*$ the $w^*$-convex closed hull of $\{\psi(0); \eta \in M\}$.

Using the same argument as in Lemma 3(iii) we see that the elements $y$ of $A^*$ are either Bochner integrals $\int f \psi dA$ as before, or else $0 < \Phi(E) y < \epsilon < 1$, for all $E \in J$, with $\epsilon$ depending only on $y$. Here, of course, we use the fact that $w^* \text{-lim } y = 0$. The unit ball for the norm in $X^*$ is $w^* \text{-co}(B_{1/2}, \cup (A^*) \cup (-A^*) \cup \cdots)$. Since this is the unit ball of a dual norm, the proof is complete in the special case $X = Z$.

When $X$ is not separable, it contains a separable, non-reflexive subspace $Z$. The functionals $\Phi(E)$ belong to $Z^* \subseteq X^*$, but the elements $\psi(0)$ are not defined over all of $X$, that is, they are not elements of $X^*$. Let $\theta(0)$ be an extension of $\psi(0)$ to all of $X$, of the same norm. No matter how this is done, $\Phi(E)(\theta(0))$ has the same value, since $\Phi(E) \subseteq Z^*$. This allows us to define (in a somewhat arbitrary manner) a $w^*$-closed set $A^* \subseteq X^*$. These elements of $X^*$ are extensions of the functionals constructed before, and serve exactly the same purpose.

In place of this process of two stages, is it possible to perform the first step in $X^*$ in the case of certain non-reflexive spaces ("Grothendieck spaces" [7, p. 150]), $w^*$-convergent sequences in $X^*$ must converge weakly, thereby obstructing the simplification.

References