Definition 1 (Sublinear functional) $\rho : X \to \mathbb{R}$ is a sublinear functional if

$$\begin{aligned} \rho(x+y) &\leq \rho(x) + \rho(y) && x, y \in X \\ \rho(\lambda x) &= \lambda \rho(x) && \lambda > 0 \end{aligned}$$

I'm splitting the proof into two parts because the infinite-dimensional part of the extension is independent from ZF.

Lemma 2 (one-dimensional extension) *Suppose E is a subspace of X. Let $E' = E \oplus \operatorname{span}\{e\}$ be a one-dimension-larger subspace of X. If $\rho : X \to \mathbb{R}$ is a sublinear functional and $f : E \to \mathbb{R}$ is a linear functional bounded by ρ, then there is an extension f' of f to E' such that f' is also bounded by ρ.*

Proof. Proceed by finding a bound on potential extensions to E' where $f' \leq \rho$ still. Suppose we already have a suitable f'. Let $f'(e) = \alpha$. To extract a bound on α, consider the following inequalities:

$$f'(x-e) = f(x) - \alpha \leq \rho(x-e) \tag{1}$$
$$f'(y+e) = f(y) + \alpha \leq \rho(y+e) \tag{2}$$

Combining these gives

$$f(x) - \rho(x-e) \leq \alpha \leq \rho(y+e) - f(y)$$

Because x and y are arbitrary, this requires

$$\sup\{f(x) - \rho(x-e)\} \leq \alpha \leq \inf\{\rho(y+e) - f(y)\}. \tag{3}$$

It is necessary to show the left hand side of eq. (3) is, in fact, less than or equal to the right hand side. But

$$f(x) + f(y) = f(x+y) \leq \rho(x+y) \leq \rho(x-e) + \rho(y+e).$$

Rearranging gives $f(x) - \rho(x-e) \leq \rho(e+y) - f(y)$. Recalling, again, that x and y are arbitrary gives

$$\sup\{f(x) - \rho(x-e)\} \leq \inf\{\rho(e+y) - f(y)\}.$$

Therefore there is at least one such α.

Now, we need to check that satisfying eq. (3) is sufficient to give a bounded extension of f. Suppose α satisfies eq. (3) and, if $x \in E$, define the extension $f'(x+\lambda e) = f(x) + t\alpha$. If $\lambda > 0$,

$$f'(x+\lambda e) = \lambda\big(f(x/\lambda) + \alpha\big) \underbrace{\leq \lambda\rho(x/\lambda + e)}_{\text{by eq. (2)}} = \rho(x+\lambda e)$$

If $\lambda = -\mu < 0$,

$$f'(x-\mu e) = \mu\big(f(x/\mu) - \alpha\big) \overbrace{\leq \mu\rho(x/\mu - e)}^{\text{by eq. (1)}} = \rho(x+\lambda e)$$

proving sufficiency. $\square$

By induction, this proves we can extend to a finite number of dimensions. The typical approach uses Zorn's lemma to extend to the infinite dimensional case:

Theorem 3 (Hahn-Banach theorem) *If $f : E \to \mathbb{R}$ is bounded by a sublinear functional $\rho : X \to \mathbb{R}$, then f extends to a functional $f' : X \to \mathbb{R}$ which is bounded by ρ on the whole space.*

Proof. The set of linear extensions of f from subspaces of X to $\mathbb{R}$ is a poset ordered by inclusion. But if f_α are chain of extensions of f, they must agree on every subset of X on which they are defined. Hence $\sup f_\alpha := \bigcup f_\alpha$ is well-defined and gives an upper bound on the chain. By Zorn's lemma, there is a maximal f^*.

Suppose f^* is not defined on E, not the entirety of X. Then pick some e on which f^* is not defined. By lemma 2 (one-dimensional extension), there is some $f^{*\prime}$ extending f^* to $E \oplus \operatorname{span}\{e\}$. This contradicts maximality; hence, $f^* : X \to \mathbb{R}$. □

Here is a sketch of an alternate proof using the compactness of first order logic, which is strictly weaker than the axiom of choice.

Proof. We need to axiomatize the statement of the theorem. Consider the domain of discourse of points of the form $X \times \mathbb{R}$. Let each needed $X \times \mathbb{R}$ be a constant in this language. Write the axioms of a vector space and a linear map in terms of the domain of discourse. Take the predicate $\square \in E$ for every subspace E. Then consider the axiom schema $\forall (x, y) x \in E \implies y \leq \rho(x)$ encoding "bounded by ρ". Inductively applying lemma 2 (one-dimensional extension) proves that each finite subtheory has a model. By compactness, there must be a model of the whole theory, which is a function defined on all of X, and this model is the desired extension. □

Although Hahn-Banach is weaker than logical compactness, it is still strong enough to construct the Banach-Tarski paradox.