Before getting to the proof of the result below, I want emphasize the relation of the hypothesis

"Given any sequence of points, $\{p_n\}$, $p_n \in S$, there is at least one $p_\infty \in S$ such that for every $r > 0$, $p_n \in B_r(p_\infty)$ for *infinitely many* n."

to the more standard hypothesis

"Given any sequence of points, $\{p_n\}$, $p_n \in S$, there is at least one $p_\infty \in S$ and a subsequence[1] of $\{p_n\}$, $\{p_{n_k}\}$, such that for every $r > 0$ there is a K such that $p_{n_k} \in B_r(p_\infty)$ when $k \geq K$". This is usually shortened to "every sequence of points in S has a convergent subsequence with its limit in S."

That the second hypothesis implies the first should be clear. The two hypotheses are actually equivalent. To see that the first implies the second argue this way: The first hypothesis implies that there is an n, n_1, such that $p_{n_1} \in B_1(p_\infty)$. Since $p_n \in B_{1/2}(p_\infty)$ for infinitely many n, there is an n, n_2, such that $p_{n_2} \in B_{1/2}(p_\infty)$ *and* $n_2 > n_1$. Continue this way, using the radii $1/3$, $1/4$, ... , $1/k$, ... , and choosing each n_k greater than the one before it. This way you get a subsequence $\{p_{n_k}\}_{k=1}^{\infty}$ such that $p_{n_k} \in B_{1/k}(p_\infty)$ for all k. So the second hypothesis holds – just take K to be the first integer such that $K > 1/r$.

Bolzano-Weierstrass Compactness implies Heine-Borel Compactness

This is the difficult direction in the proof that, even in the generality of metric spaces, the Bolzano-Weierstrass definition of compact sets is equivalent to the Heine-Borel definition of compact sets.

Step 1: $\mathcal{X}$ is B-W compact implies that $\mathcal{X}$ is totally bounded.

Proof: Suppose not. Then there is an n_0 such that no finite set of open balls of radius $1/n_0$ covers $\mathcal{X}$. So, picking $x_1 \in \mathcal{X}$ at random, one can define a sequence in $\mathcal{X}$ by choosing

$$x_{m+1} \in (\cup_{i=1}^{m} B(x_i, 1/n_0))^c.$$

Since $d(x_{m+1}, x_i) \geq 1/n_0,\ i = 1, .., m$, this sequence has no convergent subsequences.

The remainder of the argument is based on the following. For any nonempty $S \subset \mathcal{X}$ define $d(x, S) = \inf_{z \in S} d(x, z)$. Then it follows that (this is good exercise)

$$d(x, S) \leq d(x, y) + d(y, S). \tag{1}$$

We suppose that $\mathcal{X}$ is B-W compact and $\{\mathcal{O}_\alpha,\ \alpha \in A\}$ is an open cover of $\mathcal{X}$. Thus we need to show that $\{\mathcal{O}_\alpha,\ \alpha \in A\}$, contains a finite subcover.

Step 2: Let $\mathcal{O}_{\alpha,n} = \{x \in \mathcal{O}_\alpha : d(x, \mathcal{O}_\alpha^c) > 1/n\}$. If $\mathcal{X} \subset \cup_{\alpha \in A} \mathcal{O}_{\alpha, n_0}$ for some n_0, then $\{\mathcal{O}_\alpha,\ \alpha \in A\}$, contains a finite subcover.

Proof: By total boundedness there are $x_i,\ i = 1, .., N$, such that

[1] To be a subsequence of the sequence $\{p_n\}_{n=1}^{\infty}$ the subset $\{p_{n_k}\}_{k=1}^{\infty}$ just needs to satisfy $n_k < n_{k+1}$ for all k. Note that this implies $n_k \geq k$. This often useful.

$\mathcal{X} \subset \cup_{i=1}^N B(x_i, 1/n_0)$, and by the hypothesis here $x_i \in \mathcal{O}_{\alpha_i, n_0}$ for some choice of $\alpha_i,\ i = 1, .., N$. By 1) if $y \in B(x_i, 1/n_0)$, then $d(y, \mathcal{O}^c_{\alpha_i}) \geq d(x_i, \mathcal{O}^c_{\alpha_i}) - d(y, x_i) > 0$. Thus $B(x_i, 1/n_0) \subset \mathcal{O}_{\alpha_i}$ and $\mathcal{X} \subset \cup_{i=1}^N \mathcal{O}_{\alpha_i}$.

Step 3: If there is no n such that $\{\mathcal{O}_{\alpha,n},\ \alpha \in A\}$ covers $\mathcal{X}$, then we can choose a sequence $\{x_n\}$ such that, for all n, $x_n \in (\cup_{\alpha \in A} \mathcal{O}_{\alpha,n})^c$. If $x_{n_k} \to x_\infty$ as $k \to \infty$, then $x_\infty \in \mathcal{O}_{\alpha_0}$ for some α_0 and, hence $x_\infty \in \mathcal{O}_{\alpha_0, n_0}$ for some n_0, because $\mathcal{O}_{\alpha_0}$ is open. Note that $d(x_n, \mathcal{O}^c_{\alpha_0}) \leq 1/(2n_0)$, when $n \geq 2n_0$ because $x_n \notin \mathcal{O}_{\alpha_0, 2n_0}$. So, when $n \geq 2n_0$,

$$d(x_n, x_\infty) \geq d(x_\infty, \mathcal{O}^c_{\alpha_0}) - d(x_n, \mathcal{O}^c_{\alpha_0}) \geq 1/n_0 - 1/(2n_0) = 1/(2n_0).$$

This contradicts $x_{n_k} \to x_\infty$ as $k \to \infty$, and that completes the proof.