Test 1: Take Home

Huan Q. Bui

1. Let X denote the set of all irrational numbers x with $\sqrt{2} \leq x \leq 2\sqrt{2}$, and with the usual metric $d(x, y) = |x - y|$. Prove that X is not compact.

2. Let (X, d) denote any metric space. The metric space X is called "totally bounded" when, for every $\epsilon > 0$, there exists finitely many neighborhoods $N_\epsilon(x_i)$ $(i = 1, \dots n)$ such that $X \subseteq \cup_{i=1}^n N_\epsilon(x_i)$. The metric space is "bounded" when $\{d(x, y) | x, y \in X\}$ is a bounded subset of $\mathbb{R}$.

   (a) Give an example of a bounded metric space that is not totally bounded.

   (b) Prove that every totally bounded metric space is bounded

   (c) Prove that a metric space is compact if and only if it is both complete and totally bounded.

3. Let $\mathbb{R}^n$ denote the usual n-dimensional Euclidean space, with its Euclidean norm

$$\|x\| = \sqrt{\sum_{i=1}^n |x_i|^2}$$

   and corresponding metric $d(x, y) = \left\|x - y\right\|$, with $x, y \in \mathbb{R}^n$. Given an $n \times n$ matrix T, define

$$\|T\| \equiv \sup\{\|Tx\| : \|x\| \leq 1\}.$$

   (a) Prove that, for all $n \times n$ matrices X and Y, that $\|XY\| \leq \|X\|\|Y\|$.

   (b) Prove that
$$\|T\| = \inf\{M \in \mathbb{R} : \|Tx\| \leq M\|x\| \text{ for all } x \in \mathbb{R}^n\}.$$

   (c) With $x \in \mathbb{R}^n$, find $\|C_x\|$ when C_x is the $n \times n$ matrix with the coordinates of x in the first column and zeros elsewhere.

   (d) With $x \in \mathbb{R}^n$, find $\|D_x\|$ when D_x is the $n \times n$ diagonal matrix with the coordinates of x on the main diagonal, and zeros elsewhere.

   (e) With $x \in \mathbb{R}^n$, find $\|R_x\|$ when R_x is the $n \times n$ matrix with the coordinates of x in the first row and zeros elsewhere.

4. Let T be an $n \times n$ matrix, with $\|T\|$ defined as in the previous problem. Prove that

$$\inf\{\|T^m\|^{\frac{1}{m}} : m \in \mathbb{N}\} = \sup\{|\alpha| : \alpha \text{ an eigenvalue of } T\}$$

Test 1: Solution

1. Let X denote the set of all irrational numbers x with $\sqrt{2} \leq x \leq 2\sqrt{2}$, and with the usual metric $d(x, y) = |x - y|$. Prove that X is not compact.

Proof: Since $X \subset \mathbb{R}$, it suffices to show X is either not bounded or not closed (or neither). X is evidently bounded, so we will show X is not closed. To this end, we claim X^c is not open, where

$$X^c = \underbrace{\left(\mathbb{R} \setminus [\sqrt{2}, 2\sqrt{2}]\right)}_{A} \cup \underbrace{\left\{r \in \mathbb{Q} | \sqrt{2} < r < 2\sqrt{2}\right\}}_{B}.$$

We note that $A \cap B = \emptyset$ and let $\epsilon > 0$ be given. Consider $r \in B \subset X^c$ and $\mathcal{N}_\epsilon(r)$. We want to show that $\mathcal{N}_\epsilon(r) \not\subset X^c$, i.e., $\exists x \in X$ such that $x \in \mathcal{N}_\epsilon(r)$.

Because $\mathbb{Q}$ is dense in $\mathbb{R}$, $\exists r' \in B$ such that $r' \in \mathcal{N}_\epsilon(r)$. Without loss of generality, suppose $r' < r$. Let an irrational number $\bar{x}$ be given. By the denseness of $\mathbb{Q}$, there is a rational number $q \in (r'/\bar{x}, r/\bar{x})$ such that $\bar{x}q \in (r', r)$, hence contained in $\mathcal{N}_\epsilon(r)$. Call $x = \bar{x}q$. Since x is a product of an irrational number and a rational number, x is irrational, hence $x \notin B \subset X^c$. Because $\mathcal{N}_\epsilon(r) \not\subset B \subset X^c$ and $A \cap B = \emptyset$, $\mathcal{N}_\epsilon(r) \not\subset X^c$. So, X^c is not open $\iff$ X is not closed, which implies X is not compact.

$\square$

2. Let (X, d) denote any metric space. The metric space X is called "totally bounded" when, for every $\epsilon > 0$, there exists finitely many neighborhoods $N_\epsilon(x_i)$ $(i = 1, \dots n)$ such that $X \subseteq \cup_{i=1}^n N_\epsilon(x_i)$. The metric space is "bounded" when $\{d(x, y) : x, y \in X\}$ is a bounded subset of $\mathbb{R}$.

1. Give an example of a bounded metric space that is not totally bounded.
2. Prove that every totally bounded metric space is bounded
3. Prove that a metric space is compact if and only if it is both complete and totally bounded.

---

1. Consider $X = [0, 1]$ with the metric:

$$d(x, y) = \begin{cases} 1, & x \neq y \\ 0, & x = y \end{cases}$$

   By Problem 10, Chapter 2, Baby Rudin, (X, d) is a metric space. Clearly X is bounded because $X \subset \mathcal{N}_{r=2}(0)$. However, X is not totally bounded. Set $\epsilon = 1/2$, then for any x, $\mathcal{N}_\epsilon(x) = \{x\}$. It follows that for any finite set $\{x_1, \dots, x_n\}$,

$$\cup_{i=1}^n \mathcal{N}_\epsilon(x_i) = \{x_1, \dots, x_n\} \not\supseteq [0, 1].$$

   □

2. Let a totally bounded metric space (X, d) be given. By definition, $\forall \epsilon > 0, \exists n \in \mathbb{N}$ such that $X \subseteq \cup_{i=1}^n \mathcal{N}_\epsilon(x_i)$. Let $\epsilon > 0$ be given. Consider the points a, b in X where $a \in \mathcal{N}_\epsilon(x_i)$ and $b \in \mathcal{N}_\epsilon(x_j)$. Then we have

$$d(a, b) \leq d(a, x_i) + d(x_i, x_j) + d(x_j, b) < \epsilon + d(x_i, x_j) + \epsilon.$$

   Since there are only finitely many values of $d(x_i, x_j), 0 \leq d(a, b) < 2\epsilon + \sup\{d(x_i, x_j) | i, j = 1, \dots, n\}$. Thus, $\{d(a, b) | a, b \in X\}$ is a bounded subset of $\mathbb{R}$, which implies (X, d) is bounded.

   □

3. $(\rightarrow)$ Let a metric space (X, d) be given. Suppose (X, d) is compact, i.e., each of its open cover has a finite subcover. We want to show (X, d) is complete and totally bounded.

   - (Completeness) To prove: Every Cauchy sequence in X converges.

     Let a Cauchy sequence $\{x_n\} \subset X$ be given.

     – If the set $\Gamma \subset X$ of the terms of $\{x_n\}$ is finite then $\{x_n\}$ converges to some term $x_k \in \Gamma$, because by definition $x_i, x_j \in \{x_n\}$ get arbitrarily close for sufficiently large i, j.

– If $\Gamma \subset X$ is infinite then Γ contains its limit point p because X is compact (theorem 2.37, Baby Rudin). We want to show $x_n \longrightarrow p$. To this end, let $\epsilon > 0$ be given and set $\epsilon' = \epsilon/2$. Since $\{x_n\}$ is Cauchy, $\exists N \in \mathbb{N}$ such that whenever $m, n \geq N$,

$$d(x_m, x_n) < \epsilon' = \frac{\epsilon}{2}. \tag{1}$$

We also know p is a limit point of Γ, so for $r = \epsilon' = \epsilon/2 > 0$, $\exists x_m \in \Gamma$ where $m \geq N$ such that $x_m \in \mathcal{N}_{\epsilon'}(p) \setminus \{p\} \neq \emptyset$, which means

$$d(x_m, p) \leq \epsilon' = \frac{\epsilon}{2}. \tag{2}$$

From (1) and (2), if $n \geq N$, we have that

$$d(x_n, p) \leq d(x_n, x_m) + d(x_m, p) < \epsilon' + \epsilon' = \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Thus, the Cauchy sequence $\{x_n\}$ in X converges to p in X, which implies X is complete.

- (Totally boundedness) To prove: $\forall \epsilon > 0, \exists n \in \mathbb{N}, n < \infty$, such that $X \subseteq \cup_{i=1}^{n} \mathcal{N}_\epsilon(x_i)$.
  Let a compact metric space (X, d) be given. Then the collection $\{\mathcal{N}_\epsilon(x) | x \in X\}$ forms an open cover for X. Since X is compact, there is a finite subcover, i.e., there are (finitely many) points $x_1, \ldots, x_n \in X$ such that

$$X = \cup_{i=1}^{n} \mathcal{N}_\epsilon(x_i).$$

  This shows X is totally bounded.

($\longleftarrow$) Let (X, d) be given. (X, d) is complete and totally bounded. To prove: (X, d) is compact.

Let the collection $\{\mathcal{N}_\epsilon\}$ be an open cover for X. Assume (to get a contradiction) that $\{\mathcal{N}_\epsilon\}$ has no finite subcover for X. Let $\alpha = \text{diam}(X)$, which exists because X is (totally) bounded. Since X is totally bounded, X can be covered by finitely many closed ball $\mathcal{B}_{\alpha/4}(x_i)$ with $x_i \in X$. It follows from our assumption that at least one $\mathcal{B}_{\alpha/4}(x_j)$ intersected with X cannot be finitely covered by $\{\mathcal{N}_\epsilon\}$. Call $X_1 = \mathcal{B}_{\alpha/4}(x_j) \cap X$, then X_1 is a closed subset of X with $\text{diam}(X_1) \leq \alpha/2$ (X_1 closed by theorems 2.24(b) and 2.34, Baby Rudin). Repeating this argument gives us a nested sequence of closed sets $X_n \subset X$ with $\text{diam}(X_n) \leq \alpha/2^n$ where each X_n cannot be finitely covered by $\{\mathcal{N}_\epsilon\}$.

Now, for each n, consider $x_n \in X_n$. Then $\{x_n\}$ is Cauchy, by the construction of the closed subsets X_n. Because X is complete, $\{x_n\}$ converges with some limit $p \in X$. Since each X_n is closed, we have that $p \in \cap_{n=1}^{\infty} X_n$. Further, because $\text{diam}(X_n) \longrightarrow 0$ as

$n \to \infty$, we must have that $\cap_{n=1}^{\infty} A_n = \{p\}$, the set with a single element p. Consider any $\mathcal{N} \in \{\mathcal{N}_\epsilon\}$ with $p \in \mathcal{N}$. $\mathcal{N}$ is open, so there exists $r > 0$ such that $\mathcal{N}_r(p) \subset \mathcal{N}$. Take $n \in \mathbb{N}$ such that $d(p, x_n) < r/2$ and $\text{diam}(X_n) < r/2$, then $X_n \subset \mathcal{N}_r(p) \subset \mathcal{N} \in \{\mathcal{N}_\epsilon\}$, which contradicts the assumption that X_n cannot be finitely covered by $\{\mathcal{N}_\epsilon\}$. So, $\{\mathcal{N}_\epsilon\}$ has a finite subcover for X, so (X, d) is compact.

□

Reference

For part 3. of this problem, I used the approach given by Anton R. Schep, which is presented in *Compact sets in metric spaces, Notes for Math 703*, link here. I also found different versions of this proof which show compactness via the convergence subsequences, but I like Schep's approach best because it uses the definition of compactness.

3. Let $\mathbb{R}^n$ denote the usual n-dimensional Euclidean space, with its Euclidean norm

$$\|x\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

and corresponding metric $d(x, y) = \left\|x - y\right\|$, with $x, y \in \mathbb{R}^n$. Given an $n \times n$ matrix T, define

$$\|T\| \equiv \sup\{\|Tx\| : \|x\| \leq 1\}.$$

1. Prove that, for all $n \times n$ matrices X and Y, that $\|XY\| \leq \|X\|\|Y\|$.
2. Prove that
$$\|T\| = \inf\{M \in \mathbb{R} : \|Tx\| \leq M\|x\| \text{ for all } x \in \mathbb{R}^n\}.$$
3. With $x \in \mathbb{R}^n$, find $\|C_x\|$ when C_x is the $n \times n$ matrix with the coordinates of x in the first column and zeros elsewhere.
4. With $x \in \mathbb{R}^n$, find $\|D_x\|$ when D_x is the $n \times n$ diagonal matrix with the coordinates of x on the main diagonal, and zeros elsewhere.
5. With $x \in \mathbb{R}^n$, find $\|R_x\|$ when R_x is the $n \times n$ matrix with the coordinates of x in the first row and zeros elsewhere.

---

1. <u>To prove</u>: $\|XY\| \leq \|X\|\|Y\|$.

   We first show that $\|Yx\| \leq \|Y\|\|x\|$. Suppose (to get a contradiction) that $\|Yx\| > \|Y\|\|x\|$, then it follows that

   $$\frac{1}{\|x\|}\|Yx\| > \|Y\| \implies \left\|Y\frac{x}{\|x\|}\right\| > \|Y\|.$$

   Because $x/\|x\|$ is a unit vector, this contradicts the definition of $\|Y\|$. Thus, $\|Yx\| \leq \|Y\|\|x\|$. It follows that

   $$\begin{aligned}
   \|XY\| &= \sup\{\|XYx\| : \|x\| \leq 1\} \\
   &\leq \sup\{\|X\|\|Yx\| : \|x\| \leq 1\} \\
   &= \|X\| \sup\{\|Yx\| : \|x\| \leq 1\} \\
   &= \|X\|\|Y\|
   \end{aligned}$$

   □

2. <u>To prove</u>: $\sup\{\|Tx\| : \|x\| \leq 1\} = \inf\{M \in \mathbb{R} : \|Tx\| \leq M\|x\| \forall x \in \mathbb{R}^n\}$.

   Let

   $$\begin{aligned}
   a &= \inf\{M \in \mathbb{R} : \|Tx\| \leq M\|x\| \forall x \in \mathbb{R}^n\} \\
   b &= \sup\{\|Tx\| : \|x\| \leq 1\}
   \end{aligned}$$

   We want to show $a \leq b$ and $b \leq a$.

- By definition, $\|Tx\| \le a\|x\| \, \forall x \in \mathbb{R}^n$. In particular, this holds for $\|x\| \le 1$. And so, $b \ge \|Tx\| \le a\|x\| \le a$, i.e., $b \le a$.
- Consider the quantity

$$c = \sup \left\{ \frac{\|Tx\|}{\|x\|} : x \in \mathbb{R}^n \setminus \{0\} \right\}.$$

Clearly, $\|Tx\| \le d\|x\|$ for all nonzero $x \in \mathbb{R}^n$. So, $a \le c$, by the definition of a. Consider another quantity:

$$d = \sup\{\|Tx\| : \|x\| = 1\}.$$

For any nonzero $x \in \mathbb{R}^n$, $x/\|x\|$ is a unit vector, which means $\|Tx\|/\|x\| = \|T(x/\|x\|)\| \le d$. By the definition of c, we have that $c \le d$ and thus $a \le c \le d$. Finally, $d \le b$ clearly because d is a supremum taken over fewer terms than b.

Thus, $a \le c \le d \le b \le a$, which implies $a = b$.

□

3. Let $x = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}^\top \in \mathbb{R}^n$ be given. Then C_x has the form

$$C_x = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ x_n & 0 & \dots & 0 \end{pmatrix}_{n \times n}.$$

Let $y = \begin{pmatrix} y_1 & \dots & y_n \end{pmatrix}^\top \in \mathbb{R}^n$ be given, then clearly $C_x y = y_1 x \implies \left\|C_x y\right\| = |y_1|\|x\|$. By definition,

$$\begin{aligned} \|C_x\| &= \sup\left\{\left\|C_x y\right\| : \left\|y\right\| \le 1\right\} \\ &= \sup\left\{|y_1|\|x\| : \left\|y\right\| \le 1\right\} \\ &= \|x\| \sup\left\{|y_1| : \left\|y\right\| \le 1\right\} \\ &= \|x\|, \quad \text{attained when taking } y = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}^\top. \end{aligned}$$

Thus, $\|C_x\| = \|x\|$. □

4. Let $x = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}^\top \in \mathbb{R}^n$ be given. Then D_x has the form

$$D_x = \operatorname{diag}(x_1, \dots, x_n).$$

Let $y = \begin{pmatrix} y_1 & \dots & y_n \end{pmatrix}^\top \in \mathbb{R}^n$ be given, then clearly

$$\left\|D_x y\right\| = \left\|\begin{pmatrix} x_1 y_1 & \dots x_n y_n \end{pmatrix}^\top\right\| = \sqrt{\sum_{i=1}^n |x_i y_i|^2} = \sqrt{\sum_{i=1}^n x_i^2 y_i^2}.$$

By definition,

$$\begin{aligned}\|D_x\| &= \sup\left\{\left\|D_x y\right\| : \left\|y\right\| \leq 1\right\} \\ &= \sup\left\{\left\|D_x y\right\| : \left\|y\right\| = 1\right\}\end{aligned}$$

where we have used the previous result: $a \leq c \leq d \leq b \leq a$ in the second equality. With this,

$$\begin{aligned}\|D_x\| &= \sup\left\{\sqrt{\sum_{i=1}^{n} x_i^2 y_i^2} : \left\|y\right\| = 1\right\} \\ &\leq \sup\left\{\sqrt{\sum_{i=1}^{n}\left(\max_{1\leq i\leq n}|x_i|\right)^2 y_i^2} : \left\|y\right\| = 1\right\} \\ &= \sup\left\{\max_{1\leq i\leq n}|x_i|\sqrt{\sum_{i=1}^{n} y_i^2} : \left\|y\right\| = 1\right\} \\ &= \max_{1\leq i\leq n}|x_i| \cdot \underbrace{\sup_{\left\|y\right\|=1}\left\|y\right\|}_{1} \\ &= \max_{1\leq i\leq n}|x_i|,\end{aligned}$$

with equality occurring when $y = e_{(m(i))}$ where $e_{(j)}$ is one of the standard basis vectors with 1 at the jth coordinate and zero elsewhere, and $m(i)$ is the index of the largest coordinate (in magnitude) of x. In other words, $||D_x||$ is the absolute value of the largest coordinate of x (in magnitude). Thus, $||D_x|| = \max_{1\leq i\leq n}|x_i|$. $\square$

5. Let $x = \begin{pmatrix} x_1 & \ldots & x_n\end{pmatrix}^\top \in \mathbb{R}^n$ be given. Then C_x has the form

$$R_x = \begin{pmatrix} x_1 & \ldots & x_n \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 0\end{pmatrix}_{n\times n}.$$

Let $y = \begin{pmatrix} y_1 & \ldots & y_n\end{pmatrix}^\top \in \mathbb{R}^n$ be given, then clearly,

$$\left\|R_x y\right\| = \left\|\begin{pmatrix}\sum_{i=1}^n x_i y_i & 0 & \ldots & 0\end{pmatrix}^\top\right\| = \left\|\sum_{i=1}^{n} x_i y_i \begin{pmatrix}1 & 0 & \ldots & 0\end{pmatrix}^\top\right\| = \left|\sum_{i=1}^{n} x_i y_i\right|.$$

By definition,

$$
\begin{aligned}
\|R_x\| &= \sup\left\{\left\|R_x y\right\| : \left\|y\right\| \le 1\right\} \\
&= \sup\left\{\left\|R_x y\right\| : \left\|y\right\| = 1\right\} \\
&= \sup\left\{\left\|R_x y\right\| : \left\|y\right\| = 1\right\} \\
&= \sup\left\{\left|\sum_{i=1}^{n} x_i y_i\right| : \left\|y\right\| = 1\right\} \\
&\le \sup\left\{\sqrt{\sum_{i=1}^{n} x_i^2 \sum_{j=1}^{n} y_j^2} : \left\|y\right\| = 1\right\}, \quad \text{Cauchy-Schwartz} \\
&= \|x\|,
\end{aligned}
$$

where equality occurs if and only if y is a multiple of x, under the constraint $\left\|y\right\| = 1$. This means equality is attained if and only if $y = x/\|x\|$. Thus, $\|R_x\| = \|x\|$. $\square$

Reference

For Part 2. of this problem, I referred to Proposition 2.1, Chapter III: Banach Spaces, in John Conway's *A Course in Functional Analysis*, 2nd Edition, to define the quantities c, d for the proof.

4. Let T be an $n \times n$ matrix, with $||T||$ defined as in the previous problem. Prove that

$$\inf\{||T^m||^{\frac{1}{m}} : m \in \mathbb{N}\} = \sup\{|\alpha| : \alpha \text{ an eigenvalue of } T\}.$$

Note to Ben: the proof below is a combination of Internet/book search and my notes from Prof. Livshits's MA353: Matrix Analysis from S'19. The statement of the problem is similar to the statement of the Beurling-Gelfand spectral radius theorem. However, the proof found in Rudin's *Functional Analysis*, section 10.13, is too advanced for me. I found another approach by Joel E. Tropp (Prof. of Mathematics at Caltech), here, which uses Jordan canonical form (which I learned in MA353) and the fact that all norms on a finite-dimensional vector space are equivalent (which I learned from Prof. Randles) to prove the above statement. However, instead of showing the statement holds for the ∞-norm like Joel E. Tropp did, I will be using the $||\cdot||_{\text{HS}}$ norm, since I have done this in MA353.

Before getting to the proof, I want to give a lemma which is useful later in the proof.

Lemma 4.1. Suppose that $\{x_{1_n}\}, \{x_{2_n}\}, \ldots, \{x_{k_n}\}$ are sequences of positive numbers such that $\{(x_{i_n})^{1/n}\} \to \alpha_i$ for each $i = 1, 2, \ldots, k$. Then

$$\left\{(\alpha_1^n + \alpha_2^n + \cdots + \alpha_k^n)^{1/n}\right\} \to \sup_i\{\alpha_i\}.$$

It follows that

$$\left\{(x_{1_n} + x_{2_n} + \cdots + x_{k_n})^{1/n}\right\} \to \sup_i\{\alpha_i\}.$$

Proof of Lemma 4.1.: We assume (without loss of generality) that $\sup_i \alpha_i = \alpha_1$. Then, any α_i can be written as $\delta_i\alpha_1$ where δ_i is some positive number less than or equal to 1. It follows that

$$(\alpha_1^n + \alpha_2^n + \cdots + \alpha_k^n)^{1/n} = \alpha_1(1 + \delta_2^n + \cdots + \delta_k^n)^{1/n}.$$

The number $(1 + \delta_2^n + \cdots + \delta_k^n)$ is at most k. Thus, when $n \to \infty$, $(1 + \delta_2^n + \cdots + \delta_k^n)$ tends to 1. Therefore, $\lim\limits_{n\to\infty}(\alpha_1^n + \alpha_2^n + \cdots + \alpha_k^n)^{1/n} = \alpha_1$, i.e., $\left\{(\alpha_1^n + \alpha_2^n + \cdots + \alpha_k^n)^{1/n}\right\} \to \sup_i\{\alpha_i\}$. Since $\{(x_{i_n})^{1/n}\} \to \alpha_i$ for each $i = 1, 2, \ldots, k$, it follows that $\left\{(x_{1_n} + x_{2_n} + \cdots + x_{k_n})^{1/n}\right\} \to \sup_i\{\alpha_i\}$. Δ

<u>*Proof of problem statement*</u>:

I will use (without proving) the fact that the Hilbert-Schmidt norm $\|\cdot\|_{\text{HS}}$ and the operator norm $\|\cdot\|$ are equivalent, i.e., there are positive numbers $a, b > 0$ such that for any $n \times n$ matrix T, $a\|T\|_{\text{HS}} \le \|T\| \le b\|T\|_{\text{HS}}$. (A general theorem about equivalence of norms on finite-dimensional vector spaces is provided by theorem 2.4-5, Erwin Kreyszig's *Introductory Functional Analysis with Applications*). The fact about "equivalence of norms" allows me to translate my result using the Hilbert-Schmidt norm to the operator norm defined in Problem 3. In other words, if I could show that the problem statement holds for the Hilbert-Schmidt norm, then I could argue that it also holds when the operator norm is used.

Let $\rho(T) = \sup\{|\alpha| : \alpha \text{ an eigenvalue of } T\}$ denote the *spectral radius* of T. For any $n \times n$ matrix T, we want to first show that

$$\rho(T) = \lim_{n\to\infty} \|T^n\|_{\mathrm{HS}}^{1/n}.$$

Any $n \times n$ matrix T can be written as a direct sum of Jordan blocks following a similarity transformation. Suppose that $\mathcal{J} = S^{-1}TS = \bigoplus_{i=1}^{s} \mathcal{J}_i$, where each $\mathcal{J}_i$ is a Jordan block. Clearly, $\rho(T) = \rho(\mathcal{J})$ because $T \sim \mathcal{J}$. Now, we want to consider the relationship between $\|T^n\|^{1/n}$ and $\|\mathcal{J}^n\|^{1/n}$:

$$\|T^n\|^{1/n} = \left\|(S^{-1}\mathcal{J}S)^n\right\|^{1/n} = \left\|S\mathcal{J}^n S^{-1}\right\|^{1/n} \leq \left(\|S\|\left\|S^{-1}\right\|\right)^{1/n} \|\mathcal{J}^n\|^{1/n}$$

and

$$\|T^n\|^{1/n} = \left\|(S^{-1}\mathcal{J}S)^n\right\|^{1/n} = \left(\frac{\left\|S^{-1}\right\|\left\|S\mathcal{J}^n S^{-1}\right\|\|S\|}{\|S\|\|S^{-1}\|}\right)^{1/n} \geq \left(\|S\|\left\|S^{-1}\right\|\right)^{-1/n} \|\mathcal{J}^n\|^{1/n}$$

where we have used results from Problem 3 and the fact that $\left\|S^{-1}\right\|\left\|S\mathcal{J}^n S^{-1}\right\|\|S\| \geq \|\mathcal{J}^n\|$ when S and S^{-1} are "absorbed" into the term in the middle. Further, in each inequality, the term $\left(\|S\|\left\|S^{-1}\right\|\right)^{\pm 1/n} \to 1$ as $n \to \infty$. Thus, it suffices to consider only the behavior of $\|\mathcal{J}^n\|^{1/n}$ rather than $\|T^n\|^{1/n}$ itself, i.e., it suffices to show

$$\rho(T) = \lim_{n\to\infty} \|\mathcal{J}^n\|_{\mathrm{HS}}^{1/n}.$$

Since $\mathcal{J}$ is block-diagonal, $\mathcal{J}^n$ is a direct sum of the powers of the Jordan blocks of T, i.e., $\mathcal{J}^n = \bigoplus_{i=1}^{s} (\mathcal{J}_i)^n$. Consider a Jordan block $\mathcal{J}_i$. Let us write $\mathcal{J}_i \equiv \mathcal{J}_{\lambda,m}$ where λ is the associated eigenvalue and m is the size of $\mathcal{J}_i$. Further, we write $\mathcal{J}_{\lambda,m} = \lambda\mathcal{I} + \mathcal{N}$ where $\mathcal{I}$ is the $m \times m$ identity matrix and $\mathcal{N}$ is a nilpotent of order m. With these, we can write $(\mathcal{J}_{\lambda,m})^n$ as a sum

$$(\mathcal{J}_{\lambda,m})^n = (\lambda\mathcal{I} + \mathcal{N})^n = \lambda^n\mathcal{I} + \binom{n}{1}\lambda^{n-1}\mathcal{N} + \ldots$$

which is truncated at the term with $\mathcal{N}^m = \mathcal{O}$, the zero matrix. Since $\mathcal{N}$ has the form

$$\mathcal{N} = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix},$$

we recognize that $(\mathcal{J}_{\lambda,m})^n$ can be written as

$$(\mathcal{J}_{\lambda,m})^n = \begin{bmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & & \binom{n}{m-1}\lambda^{n-(m-1)} \\ & \lambda^n & \ddots & \\ & & \ddots & \binom{n}{1}\lambda^{n-1} \\ & & & \lambda^n \end{bmatrix}.$$

With this, we can write the formula for the Hilbert-Schmidt norm for $(\mathcal{J}_{\lambda,m})^n$ as

$$\left\|(\mathcal{J}_{\lambda,m})^n\right\|_{\mathrm{HS}}^2 = m(|\lambda|^2)^n + (m-1)\binom{n}{1}^2(|\lambda|^2)^{(n-1)} + \cdots + \binom{n}{m-1}^2(|\lambda|^2)^{(n-(m-1))}.$$

If $|\lambda| = 0$ then $\left\|(\mathcal{J}_{\lambda,m})^n\right\|_{\mathrm{HS}} = 0$, which implies

$$\lim_{n\to\infty}\left(\left\|(\mathcal{J}_{\lambda,m})^n\right\|_2\right)^{\frac{1}{n}} = \lim_{n\to\infty} 0 = 0 = |\lambda|.$$

If $|\lambda| > 0$, by factoring out $|\lambda|^n$, we get

$$\left\|(\mathcal{J}_{\lambda,m})^n\right\|_{\mathrm{HS}} = |\lambda|^n\left(m + \frac{(m-1)\binom{n}{1}^2}{|\lambda|^2} + \cdots + \frac{\binom{n}{m-1}^2}{|\lambda|^{2(m-1)}}\right)^{\frac{1}{2}}.$$

Therefore,

$$\begin{aligned}\left(\left\|(\mathcal{J}_{\lambda,m})^n\right\|_{\mathrm{HS}}\right)^{\frac{1}{n}} &= |\lambda|\left[\left(m + \frac{(m-1)\binom{n}{1}^2}{|\lambda|^2} + \cdots + \frac{\binom{n}{m-1}^2}{|\lambda|^{2(m-1)}}\right)^{\frac{1}{2}}\right]^{\frac{1}{n}} \\ &= |\lambda|\left[\left(m + \frac{(m-1)\binom{n}{1}^2}{|\lambda|^2} + \cdots + \frac{\binom{n}{m-1}^2}{|\lambda|^{2(m-1)}}\right)^{\frac{1}{n}}\right]^{\frac{1}{2}}.\end{aligned}$$

Let

$$f(n) = m + \frac{(m-1)\binom{n}{1}^2}{|\lambda|^2} + \cdots + \frac{\binom{n}{m-1}^2}{|\lambda|^{2(m-1)}}.$$

We recognize that $f(n)$ is a polynomial in n. Using logarithms and l'Hopital's rule we find $\lim_{n\to\infty}(f(n))^{\frac{1}{n}} = 1$. Thus, $\lim_{n\to\infty}\sqrt{(f(n))^{\frac{1}{n}}} = 1$, and it follows that

$$\lim_{n\to\infty}\left(\left\|(\mathcal{J}_{\lambda,m})^n\right\|_{\mathrm{HS}}\right)^{\frac{1}{n}} = |\lambda| \cdot \lim_{n\to\infty}\sqrt{(f(n))^{\frac{1}{n}}} = |\lambda| \cdot 1 = |\lambda|.$$

Back to $\mathcal{J} = \bigoplus_{i=1}^{s} \mathcal{J}_i = \bigoplus_{i=1}^{s} \mathcal{J}_{\lambda_i, m_i}$. We wish to evaluate the limit:

$$\lim_{n\to\infty} \left(\|\mathcal{J}^n\|_{\mathrm{HS}}\right)^{1/n}.$$

We have that

$$\lim_{n\to\infty} \left(\|\mathcal{J}^n\|_{\mathrm{HS}}\right)^{\frac{1}{n}} = \lim_{n\to\infty} \sqrt[n]{\left\|\bigoplus_{i=1}^{s} \left(\mathcal{J}_{\lambda_i,m_i}\right)^n\right\|_{\mathrm{HS}}} = \lim_{n\to\infty} \sqrt{\sum_{i=1}^{s} \left(\left\|\left(\mathcal{J}_{\lambda_i,m_i}\right)^n\right\|_{\mathrm{HS}}^2\right)^{\frac{1}{n}}}.$$

From an earlier argument, we know $\lim\limits_{n\to\infty} \left(\left\|\left(\mathcal{J}_{\lambda_i,m_i}\right)^n\right\|_2\right)^{\frac{1}{n}} = |\lambda_i|$. So,

$$\lim_{n\to\infty} \left(\left\|\left(\mathcal{J}_{\lambda_i,m_i}\right)^n\right\|_{\mathrm{HS}}\right)^{\frac{2}{n}} = \lim_{n\to\infty} \left(\left(\left\|\left(\mathcal{J}_{\lambda_i,m_i}\right)^n\right\|_{\mathrm{HS}}\right)^2\right)^{\frac{1}{n}} = |\lambda_i|^2.$$

If $\left\|\left(\mathcal{J}_{\lambda_j,m_j}\right)^n\right\|_{\mathrm{HS}}$ is zero for some j, then $\lambda_j = 0$, and we can drop this term from the direct sum of operators (sum to $\mathcal{J}$). Then, we can treat the positive $\left\|\left(\mathcal{J}_{\lambda_i,m_i}\right)^n\right\|_{\mathrm{HS}}^2$'s as elements of the sequences $\left\{\left(\left\|\left(\mathcal{J}_{\lambda_i,m_i}\right)^n\right\|_{\mathrm{HS}}\right)^2\right\}$, each converging to a corresponding $|\lambda_i|^2$, $i = 1, 2, \ldots, k \le s$. Using the result from Lemma 4.1., we get

$$\lim_{n\to\infty} \left(\|\mathcal{J}^n\|_{\mathrm{HS}}\right)^{\frac{1}{n}} = \lim_{n\to\infty} \sqrt{\left(\sum_{i=1}^{s} \left\|\left(\mathcal{J}_{\lambda_i,m_i}\right)^n\right\|_{\mathrm{HS}}^2\right)^{\frac{1}{n}}} = \sqrt{\sup_i(|\lambda_i|^2)} = \sup_i(|\lambda_i|) \equiv \rho(\mathcal{J}) = \rho(T).$$

We have also argued that $\lim\limits_{n\to\infty} \left(\|\mathcal{J}^n\|_{\mathrm{HS}}\right)^{\frac{1}{n}} = \lim\limits_{n\to\infty} \left(\|T^n\|_{\mathrm{HS}}\right)^{\frac{1}{n}}$, so we have

$$\lim_{n\to\infty} \left(\|T^n\|_{\mathrm{HS}}\right)^{\frac{1}{n}} = \rho(T).$$

With this we are done with the first part of the proof. Next, we want to show

$$\lim_{n\to\infty} \left(\|T^n\|_{\mathrm{HS}}\right)^{\frac{1}{n}} = \inf\{||T^m||^{\frac{1}{m}} : m \in \mathbb{N}\}.$$

To this end, we first translate our result from using the Hilbert-Schmidt norm to using the operator norm. We do this by the equivalence of norms. Since $\|\cdot\|$ is equivalent to $\|\cdot\|_{\mathrm{HS}}$, there exist positive numbers a, b such that

$$a\|T^n\|_{\mathrm{HS}} \le \|T^n\| \le b\|T^n\|_{\mathrm{HS}}.$$

Taking the nth root of this inequality and taking the limit as $n \to \infty$, we have

$$\lim_{n\to\infty} \sqrt[n]{a}\|T^n\|_{\mathrm{HS}}^{1/n} \le \lim_{n\to\infty} \|T^n\|^{1/n} \le \lim_{n\to\infty} \sqrt[n]{b}\|T^n\|_{\mathrm{HS}}^{1/n}.$$

Of course, $\lim_{n\to\infty} \sqrt[n]{a} = \lim_{n\to\infty} \sqrt[n]{b} = 1$, so we are left with

$$\lim_{n\to\infty} \|T^n\|_{\mathrm{HS}}^{1/n} \le \lim_{n\to\infty} \|T^n\|^{1/n} \le \lim_{n\to\infty} \|T^n\|_{\mathrm{HS}}^{1/n} \implies \lim_{n\to\infty} \|T^n\|_{\mathrm{HS}}^{1/n} = \lim_{n\to\infty} \|T^n\|^{1/n} = \rho(T). \tag{3}$$

To finish the proof, we want to show

$$\lim_{n\to\infty} \|T^n\|^{1/n} = \inf\{||T^m||^{\frac{1}{m}} : m \in \mathbb{N}\}.$$

Consider an eigenvalue λ of T. $\lambda \in \sigma(T)$, the spectrum of T. By the spectral mapping theorem, $\lambda^n \in \sigma(T^n)$. Since $\|T^n\| = \sup\{M \in \mathbb{R} : \|T^n x\| \le M\|x\|, \forall x \in \mathbb{R}^n\}$ (by Problem 3), we see that $|\lambda^n| \le \|T^n\|$, which implies $|\lambda| \le \|T^n\|^{1/n}$, for all $n \in \mathbb{N}$. This means $|\lambda| \le \inf\{\|T^n\|^{1/n} : n \in \mathbb{N}\}$. Now, with $\rho(T) \equiv \sup_i(|\lambda_i|)$, we have

$$\lim_{n\to\infty} \|T^n\|^{1/n} = \rho(T) \le \inf\{\|T^n\|^{1/n} : n \in \mathbb{N}\}$$

But of course, we also have by definition

$$\inf\{\|T^n\|^{1/n} : n \in \mathbb{N}\} \le \lim_{n\to\infty} \|T^n\|^{1/n}.$$

So, as desired:

$$\lim_{n\to\infty} \|T^n\|^{1/n} = \inf\{||T^m||^{\frac{1}{m}} : m \in \mathbb{N}\} \tag{4}$$

From (3) and (4),

$$\inf\{\|T^m\|^{1/m} : m \in \mathbb{N}\} = \rho(T) \equiv \sup\{|\alpha| : \alpha \text{ an eigenvalue of } T\}.$$

□

Reference

I found the statement of the theorem in section 10.13 of Rudin's *Functional Analysis* (1991), and a less advanced approach to proving it in J.A. Tropp's *An Elementary Proof of the Spectral Radius Formula for Matrices*, link here. My proof, which I actually did as an exercise in MA353, uses Jordan canonical form (like Tropp's except I used the Hilbert-Schmidt norm). The statement about the equivalence of norms is theorem 2.4-5, Erwin Kreyszig's *Introductory Functional Analysis with Applications*.