**ETH**
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Chair for Mathematical Information Science

Prof. Dr. H. Bölcskei
Sternwartstrasse 7
CH-8092 Zürich

# Mathematics of Information

Spring semester 2022

## Solutions to Problem Set 10

### Problem 1 Metric entropy of set of sequences ☕☕

a) i) Let $\epsilon < 2^{-n}$ be arbitrary but fixed, take $K = \log_2(1/\epsilon)$, and divide the interval $I_n := [-2^{-n}, 2^{-n}]$ into $L := \lceil 2^{K-n} \rceil \geq 2$ sub-intervals of equal length, centered at the points $\theta_i = -2^{-n} + \frac{(2i-1)2^{-n}}{L}$, for $i \in \{1, 2, \dots, L\}$, and each of length $\frac{2^{1-n}}{L} \leq 2^{1-K} = 2\epsilon$. By construction, for every $\theta \in I_n$, there is hence a $j \in \{1, 2, \dots, L\}$ such that $|\theta - \theta_j| \leq \epsilon$, which proves that $A_n(\epsilon) = \{\theta_i; i \in \{1, 2, \dots, L\}\}$ is an $\epsilon$-covering.

ii) For the construction of the $2\epsilon$-packing, take the points $\theta'_i = -2^{-n} + \frac{2(i-1)2^{-n}}{L-1}$, for $i \in \{1, 2, \dots, L\}$. As for every neighboring pair $\theta'_i, \theta'_j \in I_n$, it holds that $|\theta'_i - \theta'_j| = \frac{2^{1-n}}{L-1} > 2\epsilon$, we have established that $P_n(\epsilon) = \{\theta'_i; i \in \{1, 2, \dots, L\}\}$ constitutes a $2\epsilon$-packing.We finally note that $|P_n(\epsilon)| = |A_n(\epsilon)| = \lceil 2^{K-n} \rceil$.

iii) By (b.i) we have

$$N(\epsilon; I_n, \rho_1) \leq |A_n(\epsilon)| = \lceil 2^{K-n} \rceil$$

and by (b.ii),

$$M(2\epsilon; I_n, \rho_1) \geq |P_n(\epsilon)| = \lceil 2^{K-n} \rceil.$$

Combining this with $M(2\epsilon; \mathcal{C}, \rho) \leq N(\epsilon; \mathcal{C}, \rho)$ for every compact set $\mathcal{C}$ in the metric space $(\mathcal{X}, \rho)$, yields $N(\epsilon; I_n, \rho_1) = \lceil 2^{K-n} \rceil$.

b) i) An $\epsilon$-covering of $\mathcal{C}$ can be obtained by forming the Cartesian product, across $n \in \mathbb{N}$, of the $\epsilon$-coverings of $[-2^{-n}, 2^{-n}]$ according to $A(\epsilon) = \{f : \mathbb{N} \to \mathbb{R}; f(n) \in A_n(\epsilon), \forall n \in \mathbb{N}\}$, where, for $\epsilon < 2^{-n}$, $A_n(\epsilon)$ is the $\epsilon$-covering of $[-2^{-n}, 2^{-n}]$ constructed in subproblem (b.i) and $A_n(\epsilon) = \{0\}$, for $\epsilon \geq 2^{-n}$.

ii) Fix $\epsilon \leq 1/2$ and take $K = \log_2(1/\epsilon)$. By subproblems (b.iii) and (c.i), we have

$$\begin{aligned} N(\epsilon; \mathcal{C}, \rho_2) \leq |A(\epsilon)| = \prod_{n=1}^{\infty} |A_n(\epsilon)| \leq \prod_{n=1}^{\lceil K \rceil - 1} 2^{\lceil K \rceil - n} = 2^{(\lceil K \rceil - 1)\lceil K \rceil / 2} \\ \leq \left(\frac{1}{\epsilon}\right)^{\lceil K \rceil / 2} \leq \left(\frac{1}{\epsilon}\right)^{\frac{1}{2}\log_2(1/\epsilon) + C}, \end{aligned}$$

for some $C > 0$ that is independent of $\epsilon$.

## Problem 2 Minkowski dimension ☕☕

In this solution, to avoid confusion, we write $x_i$ for the $i$-th vector in the set $\{x_1, \dots, x_N\}$ and $x_{(i)}$ for the $i$-th component of the vector $x$.

a) i) The first inequality is obtained by taking the square root in the following inequality

$$\|x\|_2^2 = \sum_{i=1}^{n} \left|x_{(i)}\right|^2 \leq \sum_{i=1}^{n} \left|x_{(i)}\right|^2 + \sum_{i,j,i\neq j} \left|x_{(i)}\right| \left|x_{(j)}\right| = \left(\sum_{i=1}^{n} \left|x_{(i)}\right|\right)^2 = \|x\|_1^2, \quad (1)$$

and the second one follows by application of the Cauchy-Schwarz inequality according to

$$\|x\|_1 = \langle x, \operatorname{sgn}(x)\rangle \overset{\text{C.S.}}{\leq} \|x\|_2 \|\operatorname{sgn}(x)\|_2 \leq \sqrt{n}\|x\|_2, \quad (2)$$

with $\operatorname{sgn}(x) \in \mathbb{R}^n$ defined as

$$\operatorname{sgn}(x)_{(i)} := \begin{cases} -1, & \text{if } x_{(i)} < 0, \\ +1, & \text{if } x_{(i)} > 0, \\ 0, & \text{if } x_{(i)} = 0. \end{cases}$$

ii) Let $\{y_1, \dots, y_N\} \subset \mathbb{R}^n$ be an $\varepsilon$-covering of $K$ with respect to the $\|\cdot\|_1$-norm. For every $y \in K$, there hence exists an index $i$, $1 \leq i \leq N$, such that $\|y - y_i\|_1 \leq \varepsilon$ and therefore

$$\|y - y_i\|_2 \overset{(1)}{\leq} \|y - y_i\|_1 \leq \varepsilon.$$

We have hence established that every $\varepsilon$-covering of $K$ with respect to the $\|\cdot\|_1$-norm is also an $\varepsilon$-covering of $K$ with respect to the $\|\cdot\|_2$-norm, which in turn implies

$$\mathcal{N}(\varepsilon; K, \|\cdot\|_2) \leq \mathcal{N}(\varepsilon; K, \|\cdot\|_1).$$

Likewise, let $\{z_1, \dots, z_N\} \subset \mathbb{R}^n$ be an $(\varepsilon/\sqrt{n})$-covering of $K$ with respect to the $\|\cdot\|_2$-norm. For every $z \in K$, there hence exists an index $i$, $1 \leq i \leq N$, such that $\|z - z_i\|_2 \leq \varepsilon/\sqrt{n}$ and therefore

$$\|z - z_i\|_1 \overset{(2)}{\leq} \sqrt{n}\|z - z_i\|_2 \leq \sqrt{n}\frac{\varepsilon}{\sqrt{n}} = \varepsilon.$$

We have hence established that every $(\varepsilon/\sqrt{n})$-covering of $K$ with respect to the $\|\cdot\|_2$-norm is also an $\varepsilon$-covering of $K$ with respect to the $\|\cdot\|_1$-norm, which in turn implies

$$\mathcal{N}(\varepsilon; K, \|\cdot\|_1) \leq \mathcal{N}(\varepsilon/\sqrt{n}; K, \|\cdot\|_2).$$

iii) First note that from $\mathcal{N}(\varepsilon; K, \|\cdot\|_2) \leq \mathcal{N}(\varepsilon; K, \|\cdot\|_1)$, one has for all $\varepsilon > 0$,

$$\frac{\log_2 \mathcal{N}(\varepsilon; K, \|\cdot\|_2)}{\log_2(1/\varepsilon)} \leq \frac{\log_2 \mathcal{N}(\varepsilon; K, \|\cdot\|_1)}{\log_2(1/\varepsilon)}.$$

Taking the limit $\varepsilon \to 0^+$ on both sides yields

$$\dim_{\|\cdot\|_2}(K) \leq \dim_{\|\cdot\|_1}(K). \quad (3)$$

Likewise, it follows from $\mathcal{N}(\varepsilon; K, \|\cdot\|_1) \leq \mathcal{N}(\varepsilon/\sqrt{n}; K, \|\cdot\|_2)$, that for all $\varepsilon > 0$,

$$\frac{\log_2 \mathcal{N}(\varepsilon; K, \|\cdot\|_1)}{\log_2(1/\varepsilon)} \leq \frac{\log_2 \mathcal{N}(\varepsilon/\sqrt{n}; K, \|\cdot\|_2)}{\log_2(1/\varepsilon)} = \frac{\log_2 \mathcal{N}(\varepsilon/\sqrt{n}; K, \|\cdot\|_2)}{\log_2(\sqrt{n}/\varepsilon) - \log_2(\sqrt{n})}.$$

Taking the limit $\varepsilon \to 0^+$ on both sides yields

$$\dim_{\|\cdot\|_1}(K) \leq \dim_{\|\cdot\|_2}(K). \tag{4}$$

Combining (3) and (4), we get the desired result

$$\dim_{\|\cdot\|_1}(K) = \dim_{\|\cdot\|_2}(K).$$

iv) We proceed as above but for general norms $\|\cdot\|$ and $\|\cdot\|'$ on $\mathbb{R}^n$. From the norm equivalence in finite dimensions, it follows that there exists a constant $C \geq 1$ such that

$$C^{-1}\|x\| \leq \|x\|' \leq C\|x\|,$$

for all $x \in \mathbb{R}^n$. Let $\{y_1, \dots, y_N\} \subset \mathbb{R}^n$ be a $(C^{-1}\varepsilon)$-covering of $K$ with respect to the $\|\cdot\|$-norm. For every $y \in K$, there hence exists an index $i$, $1 \leq i \leq N$, such that $\|y - y_i\| \leq C^{-1}\varepsilon$ and therefore

$$\|y - y_i\|' \leq C\|y - y_i\| \leq C\, C^{-1}\varepsilon = \varepsilon.$$

We have hence established that every $(C^{-1}\varepsilon)$-covering of $K$ with respect to the $\|\cdot\|$-norm is an $\varepsilon$-covering with respect to the $\|\cdot\|'$-norm, which implies

$$\mathcal{N}(\varepsilon; K, \|\cdot\|') \leq \mathcal{N}(C^{-1}\varepsilon; K, \|\cdot\|). \tag{5}$$

A similar argument with the roles of the $\|\cdot\|$-norm and the $\|\cdot\|'$-norm reversed yields

$$\mathcal{N}(C\varepsilon; K, \|\cdot\|) \leq \mathcal{N}(\varepsilon; K, \|\cdot\|'). \tag{6}$$

Combining (5) and (6) allows us to conclude that

$$\frac{\log_2 \mathcal{N}(C\varepsilon; K, \|\cdot\|)}{\log_2(1/(C\varepsilon)) + \log_2(C)} \leq \frac{\log_2 \mathcal{N}(\varepsilon; K, \|\cdot\|')}{\log_2(1/\varepsilon)} \leq \frac{\log_2 \mathcal{N}(C^{-1}\varepsilon; K, \|\cdot\|)}{\log_2(1/(C^{-1}\varepsilon)) - \log_2(C)}.$$

Taking the limit $\varepsilon \to 0^+$ yields the desired result

$$\dim_{\|\cdot\|}(K) = \dim_{\|\cdot\|'}(K).$$

b) i) We first prove the result for $R = 1$. Following the Hint, we use the relation between metric entropy and the volume ratio provided in the Handout and applied to the unit ball $\mathcal{B} = \mathcal{B}' = B_{\|\cdot\|}(0, 1)$ to obtain

$$\left(\frac{1}{\varepsilon}\right)^n \leq \mathcal{N}(\varepsilon; \mathcal{B}, \|\cdot\|) \leq \left(\frac{2}{\varepsilon} + 1\right)^n, \tag{7}$$

where we used

$$vol\left(\frac{2}{\varepsilon}\mathcal{B} + \mathcal{B}'\right) = vol\left(\left(\frac{2}{\varepsilon} + 1\right)\mathcal{B}\right) = \left(\frac{2}{\varepsilon} + 1\right)^n vol\,(\mathcal{B})\,.$$

The bounds in (7) now yield $\mathcal{N}(\varepsilon; \mathcal{B}, \|\cdot\|) \asymp \varepsilon^{-n}$, which in turn implies

$$\dim(B_{\|\cdot\|}(0,1)) = \lim_{\varepsilon \to 0^+} \frac{\log_2 \mathcal{N}(\varepsilon; \mathcal{B}, \|\cdot\|)}{\log_2(1/\varepsilon)} = n.$$

For general $R > 0$, we observe that, by scaling, according to the Hint, we have

$$\mathcal{N}(\varepsilon; B_{\|\cdot\|}(0,1), \|\cdot\|) = \mathcal{N}(R\varepsilon; B_{\|\cdot\|}(0,R), \|\cdot\|), \tag{8}$$

which yields

$$\begin{aligned}
\dim(B_{\|\cdot\|}(0,R)) &= \lim_{\varepsilon' \to 0^+} \frac{\log_2 \mathcal{N}(\varepsilon'; B_{\|\cdot\|}(0,R), \|\cdot\|)}{\log_2(1/\varepsilon')} \\
&= \lim_{\varepsilon \to 0^+} \frac{\log_2 \mathcal{N}(R\varepsilon; B_{\|\cdot\|}(0,R), \|\cdot\|)}{\log_2(1/(R\varepsilon))} \\
&= \lim_{\varepsilon \to 0^+} \frac{\log_2 \mathcal{N}(R\varepsilon; B_{\|\cdot\|}(0,R), \|\cdot\|)}{\log_2(1/\varepsilon)} \\
&\overset{(8)}{=} \lim_{\varepsilon \to 0^+} \frac{\log_2 \mathcal{N}(\varepsilon; B_{\|\cdot\|}(0,1), \|\cdot\|)}{\log_2(1/\varepsilon)} \\
&= \dim(B_{\|\cdot\|}(0,1)) = n,
\end{aligned}$$

where we took $\varepsilon' = R\varepsilon$.

ii) Take $R > 0$ large enough such that $K \subset B_\infty(0,R)$, where $B_\infty(0,R)$ is the ball with respect to the infinity norm $\|\cdot\|_\infty$ centered at the origin and of radius $R$. Such an $R$ exists as $K$ is compact. This inclusion now implies a bound on the covering number according to $\mathcal{N}(\varepsilon; K, \|\cdot\|_\infty) \le \mathcal{N}(\varepsilon; B_\infty(0,R), \|\cdot\|_\infty)$, for all $\varepsilon > 0$, and consequently also on the following ratio

$$\frac{\log_2 \mathcal{N}(\varepsilon; K, \|\cdot\|_\infty)}{\log_2(1/\varepsilon)} \le \frac{\log_2 \mathcal{N}(\varepsilon; B_\infty(0,R), \|\cdot\|_\infty)}{\log_2(1/\varepsilon)}.$$

Taking the limit as $\varepsilon \to 0^+$ yields the bound $\dim_{\|\cdot\|_\infty}(K) \le \dim_{\|\cdot\|_\infty}(B_\infty(0,R))$. The desired bound according to

$$\dim(K) \le \dim(B_\infty(0,R)) \overset{(b)(i)}{=} n,$$

is now a consequence of the result in (a)(iv).

iii) Consider $K = \{x\}$, for $x \in \mathbb{R}^n$. For every $\varepsilon > 0$, we have $\mathcal{N}(\varepsilon; K, \|\cdot\|_\infty) = 1$, which yields $\dim(K) = 0 < n$.

## Problem 3 Metric entropy scaling of Lipschitz functions☕☕

a) $\Rightarrow$:
Assume $\{x_1, \dots, x_n\}$ is an $\epsilon$-covering of $X$. Let $y \in Y$ and note, that by definition of an $\epsilon$-covering, there exists $j \in \{1, \dots, n\}$ such that $\rho_X(i^{-1}(y), x_j) \le \epsilon$. By the isometry property of $i$, we have $\rho_Y(y, i(x_j)) = \rho_X(i^{-1}(y), x_j) \le \epsilon$. Consequently $\{i(x_1), \dots, i(x_n)\}$ is an $\epsilon$-covering of $Y$.
$\Leftarrow$:
Assume $\{i(x_1), \dots, i(x_n)\}$ is an $\epsilon$-covering of $Y$. Let $x \in X$ and note, that by definition of an $\epsilon$-covering, there exists $j \in \{1, \dots, n\}$ such that $\rho_Y(i(x), i(x_j)) \le \epsilon$. By the isometry property of $i$, we have $\rho_Y(x, x_j) = \rho_Y(i(x), i(x_j)) \le \epsilon$. Consequently $\{x_1, \dots, x_n\}$ is an $\epsilon$-covering of $X$.

As the covering number is the cardinality of a maximal $\epsilon$-covering, the above ensures that $N(\epsilon; X, \rho_X) = N(\epsilon; Y, \rho_Y)$.

b)
- Well-definedness:
  Let $f \in \mathcal{F}_L[a,b]$. Then $[i(f)](0) = f(a) = 0$ and for $x, x' \in [0, \Delta]$ it holds that

  $$|[i(f)](x) - [i(f)](x')| = |f(\tfrac{x}{L} + a) - f(\tfrac{x'}{L} + a)| \leq L|(\tfrac{x}{L} + a) - (\tfrac{x}{L} + a)| = |x - x'|.$$

  This shows that $i(f) \in \mathcal{F}_1[0, \Delta]$ for $f \in \mathcal{F}_L[a,b]$.
- Isometry:
  Let $f, g \in \mathcal{F}_L[a,b]$. Then, using the substitution $t := \frac{x}{L} + a$, we get that

  $$\begin{aligned}\rho_\infty(i(f), i(g)) = \|i(f) - i(g)\|_\infty &= \sup_{x\in[0,\Delta]} |f(\tfrac{x}{L} + a) - g(\tfrac{x}{L} + a)| \\ &= \sup_{t\in[a,b]} |f(t) - g(t)| = \rho_\infty(f, g).\end{aligned}$$

- Surjectivity:
  Let $F \in \mathcal{F}_1[0, \Delta]$. Then $f(x) := F(L(x - a))$ satisfies $f \in \mathcal{F}_L[a,b]$ and $i(f) = F$.
- Injectivity:
  Let $f, g \in \mathcal{F}_L[a,b]$ with $i(f) = i(g)$. Isometry implies $\rho_\infty(f,g) = \rho_\infty(i(f), i(g)) = 0$, which in turn implies $f = g$ by the definition of a metric. (This argument works in general, i.e. any isometry must be injective.)

c) This is a direct result of applying Lemma 1 with $X = \mathcal{F}_L[a,b]$, $Y = \mathcal{F}_1[0,\Delta]$, $\rho_X = \rho_Y = \rho_\infty$ and the isometric isomorphism $i$ as defined in *b*).

d) An $\epsilon$-ball (w.r.t. the metric $\rho_\infty$) around a point $\varphi \in \mathcal{F}_1[0, \Delta]$ is given by

$$B_\epsilon(\phi) = \{f \in \mathcal{F}_L[a,b] \colon \phi(x) - \epsilon \leq f(x) \leq \phi(x) + \epsilon \text{ for all } x \in [0, \Delta]\}.$$

First note that $f \in B_\epsilon(\varphi)$ if $\varphi(t_k) - \epsilon \leq f(t_k) \leq \varphi(t_k) + \epsilon$ for all[1] $k \in \{1, \dots, n\}$. This holds since a 1-Lipschitz function $f$ cannot leave and reenter the $\epsilon$-corridor around $\varphi$ on an interval $[t_{k-1}, t_k]$, since $\varphi$ changes with the maximal possible rate of 1.
Next we show by induction that for every $k \in \{1, \dots, n\}$ there exists a $\phi \in S_\epsilon$ such that $|\phi(t_k) - f(t_k)| \leq \epsilon$.

- Base case $k = 1$: Since $f(0) = 0$ and $f$ is 1-Lipschitz we have $|f(t_1) - \phi(t_1)| = |f(t_1)| \leq |t_1| = \epsilon$, which establishes the base case (as $\phi(t_1) = 0$ for all $\phi \in S_\epsilon$ by construction).
- Induction step $k \to k+1$: By the induction assumption there exists $\phi \in S_\epsilon$ with $|f(t_j) - \phi(t_j)| \leq \epsilon$ for all $j \in \{1, \dots, k\}$. The definition of $S_\epsilon$ ensures that there exist $\phi_+, \phi_- \in S_\epsilon$ with $\phi_+(x) = \phi_-(x) = \phi(x)$ for all $x \in [0, t_k]$ and $\phi_+(t_{k+1}) = \phi(t_k) + \epsilon$, $\phi_-(t_{k+1}) = \phi(t_k) - \epsilon$. As $f$ is 1-Lipschitz we have

  $$|f(t_{k+1}) - \phi(t_k)| \leq |f(t_{k+1}) - f(t_k)| + |f(t_k) - \phi(t_k)| \leq |t_{k+1} - t_k| + \epsilon = 2\epsilon.$$

  Thus we either have $|f(t_{k+1}) - \phi_+(t_{k+1})| \leq \epsilon$ or $|f(t_{k+1}) - \phi_-(t_{k+1})| \leq \epsilon$, which completes the proof by induction.

Lastly we have $|S_\epsilon| = |\{-1, 1\}^{n-1}| = 2^{n-1}$.

[1] $f(t_0) = \varphi(t_0)$ holds by definition of the set $\mathcal{F}_1[0, \Delta]$.

e) Let $\{x_1, \dots, x_n\}$ be an $\epsilon'$-covering of $X$, i.e. for every $x \in X$ there exists $i \in \{1, \dots, n\}$ such that $\rho_x(x, x_i) \leq \epsilon'$. If $\epsilon \geq \epsilon'$, the set $\{x_1, \dots, x_n\}$ clearly is also an $\epsilon$-covering. As the $\epsilon$-covering number is defined as the cardinality of a minimal $\epsilon$-covering this means we have $N(\epsilon; X, \rho) \leq N(\epsilon'; X, \rho)$ if $\epsilon \geq \epsilon'$.
Let $\epsilon > 0$ and define

$$\epsilon' := \frac{\Delta}{\lceil \frac{\Delta}{\epsilon} \rceil},$$

which is simply the largest real number smaller than $\epsilon$ for which $\Delta/\epsilon' \in \mathbb{N}$ holds. We now have

$$N(\epsilon; \mathcal{F}_1[0, \Delta], \rho_\infty) \leq N(\epsilon'; \mathcal{F}_1[0, \Delta], \rho_\infty) \leq |S_{\epsilon'}| = 2^{\frac{\Delta}{\epsilon'} - 1} = 2^{\lceil \frac{\Delta}{\epsilon} \rceil - 1} \leq 2^{\frac{\Delta}{\epsilon}}.$$

f) Let $\epsilon > 0$. Let $n' := \lceil \frac{\Delta}{\epsilon} \rceil - 1$ and $t'_k := k \frac{\Delta}{n'}$ for $k \in \{0, \dots, n'\}$. We can partition the interval $[0, \Delta]$ into the $n'$ segments

$$I'_k := [t'_{k-1}, t'_k], \quad k \in \{1, \dots, n'\}.$$

We now define

$$S'_\epsilon := \{\varphi \in C([0, \Delta]) \colon \exists (b_1, \dots, b_{n'}) \in \{-1, 1\}^{n'} \colon \forall k \in \{1, \dots, n'\} \colon (\varphi\big|_{I'_k})'(x) = b_k\},$$

which is very similar to $S_\epsilon$ in problem d), but with slightly larger intervals $I'_k$ and functions which already start going up or down on the first interval[2] $I_0$.
Let now $\phi_1, \phi_2 \in \mathcal{F}_1[0, \Delta]$. Take $t'_k$ to be the first point in $\{t'_1, \dots, t'_{n'}\}$ such that $\phi_1(t'_k) \neq \phi_2(t'_k)$. This means we have $\phi_1(t'_{k-1}) = \phi_2(t'_{k-1})$ and, by construction, $\phi_1(t'_k) = \phi_1(t'_{k-1}) + \frac{\Delta}{n'}$ as well as $\phi_2(t'_k) = \phi_2(t'_{k-1}) - \frac{\Delta}{n'}$ (or the other way around[3]). Consequently we have

$$\rho_\infty(\phi_1, \phi_2) = \sup_{x \in [0, \Delta]} |\phi_1(x) - \phi_2(x)| \geq |\phi_1(t_k) - \phi_2(t_k)| = 2\frac{\Delta}{n'} > 2\epsilon,$$

which shows that $S'_\epsilon$ is a $2\epsilon$-covering of $\mathcal{F}_1[0, \Delta]$. Thus we have $M(2\epsilon; \mathcal{F}_1[0, \Delta], \rho_\infty) \geq |S'_\epsilon| = 2^{n'} \geq 2^{\frac{\Delta}{\epsilon} - 1}$.

g) Due to $M(2\epsilon; \mathcal{F}_1[0, \Delta], \rho_\infty) \leq N(\epsilon; \mathcal{F}_1[0, \Delta], \rho_\infty)$ we get from e) and f) that

$$\frac{\Delta}{\epsilon} - 1 \leq \log_2 N(\epsilon; \mathcal{F}_1[0, \Delta], \rho_\infty) \leq \frac{\Delta}{\epsilon}.$$

Since $\Delta = L(a - b)$ and c) established that $N(\epsilon; \mathcal{F}_L[a, b], \rho_\infty) = N(\epsilon; \mathcal{F}_1[0, \Delta], \rho_\infty)$, we have proven the statement.

[2] If one would simply define $S'_\epsilon$ to be the same as $S_\epsilon$ but with larger intervals, one would get the estimate $M(2\epsilon; \mathcal{F}_1[0, \Delta], \rho_\infty) \geq 2^{\frac{\Delta}{\epsilon} - 2}$ instead, which is slightly weaker but would lead to the same end result.

[3] I.e. $\phi_1(t'_k) = \phi_1(t'_{k-1}) - \frac{\Delta}{n'}$ and $\phi_2(t'_k) = \phi_2(t'_{k-1}) + \frac{\Delta}{n'}$, which, of course, would not make any relevant difference in the argument.