Exercise 1

- One has that $\mathcal{F}$ is a Hilbert space and $\Psi : \mathcal{X} \to \mathcal{F}$, and $\forall x, x' \in \mathcal{X}$ $K(x, x') = \langle \Psi(x), \Psi(x') \rangle_{\mathcal{F}}$
So according to Aronszajn's theorem, <u>K is positive definite</u>.

- Let $\boxed{\mathcal{H} = \{ f_w , w \in \mathcal{F} \} \text{ where } \begin{array}{l} f_w : \mathcal{X} \to \mathbb{R} \\ x \mapsto \langle \Psi(x), w \rangle_{\mathcal{F}} \end{array}}$

  - One has $\mathcal{F} \neq \emptyset$ so $\mathcal{H} \neq \emptyset$
  Let $\lambda, \mu \in \mathbb{R}$ and $f, f' \in \mathcal{H}$.
  $\exists w, w' \in \mathcal{F}$, $f = f_w$, $f' = f_{w'}$
  $$\forall x \in \mathcal{X} \quad (\lambda f_w + \mu f_{w'})(x) = \lambda \langle \Psi(x), w \rangle_{\mathcal{F}} + \mu \langle \Psi(x), w' \rangle_{\mathcal{F}}$$
  $$= \langle \Psi(x), \underbrace{\lambda w + \mu w'}_{\in \mathcal{F}} \rangle_{\mathcal{F}}$$
  $$= f_{\lambda w + \mu w'} \in \mathcal{H}$$
  $\mathcal{H} \subset \mathcal{F}(\mathcal{X}, \mathbb{R})$
  So <u>$\mathcal{H}$ is a vector subspace of $\mathcal{F}(\mathcal{X}, \mathbb{R})$ so it is a $\mathbb{R}$-vector space</u>

  - Let $\boxed{\begin{array}{l} \nu : \mathcal{F} \to \mathcal{H} \\ w \mapsto f_w = \langle \Psi(\cdot), w \rangle_{\mathcal{F}} \end{array}}$

  So Ker(ν) is closed in $(\mathcal{F}, \langle \cdot, \cdot \rangle_{\mathcal{F}})$

  Let $(w_n) \in \text{Ker}(\nu)$ converging towards $w \in \mathcal{F}$. $\forall n \in \mathbb{N}$, $\forall x \in \mathcal{X}$ $\langle \Psi(x), w_n \rangle_{\mathcal{F}} = 0$; so by continuity of $\langle \cdot, \cdot \rangle_{\mathcal{F}}$ according to its second argument, $\forall x \in \mathcal{X}$, $\langle \Psi(x), w \rangle_{\mathcal{F}} = 0 \Rightarrow w \in \text{Ker}(\nu)$.
  Therefore according to the projection theorem $\mathcal{F} = (\text{Ker}\, \nu)^{\perp} \oplus \text{Ker}\, \nu$
  Let $\mathcal{F}_0 = (\text{Ker}\, \nu)^{\perp} = \{ v \in \mathcal{F}, \langle v, w \rangle_{\mathcal{F}} = 0 \ \forall w \in \text{Ker}\, \nu \}$
  and $\nu_0 = \nu_{|\mathcal{F}_0}$

↳ let $f \in \mathcal{H}$. $\exists w \in \mathcal{F}$, and thus $\exists w_0 \in \mathcal{F}_0$, such that

$$f = f_w = f_{w_0 + \underbrace{(w-w_0)}_{\in \operatorname{Ker} \mathcal{V}}} = f_{w_0} + \underbrace{f_{w-w_0}}_{=0} = f_{w_0}$$

so $\forall f \in \mathcal{H}$, $\exists w_0 \in \mathcal{F}_0$, $f = f_{w_0} = \mathcal{V}_0(w_0)$

↳ Let $w_0 \in \operatorname{Ker}(\mathcal{V}_0)$. One has $w_0 \in \operatorname{Ker} \mathcal{V} \cap (\operatorname{Ker} \mathcal{V})^\perp = \{0\}$

so $w_0 = 0$

$\Rightarrow \operatorname{Ker}(\mathcal{V}_0) = \{0\}$ ($\operatorname{Ker} \mathcal{V}_0 \subset \{0\}$ and $0 \in \operatorname{Ker} \mathcal{V}_0$)

So $\mathcal{V}_0$ is bijective.

▫ Now we define $\boxed{\begin{array}{l} \langle \cdot, \cdot \rangle_{\mathcal{H}} : \mathcal{H} \times \mathcal{H} \to \mathbb{R} \\ (f, f') \mapsto \langle \mathcal{V}_0^{-1} f, \mathcal{V}_0^{-1} f' \rangle_{\mathcal{F}} \end{array}}$

↳ Since $\langle \cdot, \cdot \rangle_{\mathcal{F}}$ is bilinear and symmetric, so is $\langle \cdot, \cdot \rangle_{\mathcal{H}}$

↳ Let $f \in \mathcal{H}$.

$\langle f, f \rangle_{\mathcal{H}} = \langle \mathcal{V}_0^{-1} f, \mathcal{V}_0^{-1} f \rangle_{\mathcal{F}} \geq 0$

and $\langle f, f \rangle_{\mathcal{H}} = 0 \Rightarrow \langle \mathcal{V}_0^{-1} f, \mathcal{V}_0^{-1} f \rangle = 0 \Rightarrow \mathcal{V}_0^{-1} f = 0 \Rightarrow f = 0$

so $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ is positive definite.

Therefore $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ defines an inner product on $\mathcal{H}$.

So $\mathcal{H}$ is a pre-Hilbert space, and it is endowed with a norm $\|f\|_{\mathcal{H}} = \sqrt{\langle f, f \rangle_{\mathcal{H}}} = \|\mathcal{V}_0^{-1} f\|_{\mathcal{F}}$

▫ Let $(f_n)_{n \geq 0}$ a Cauchy sequence in $\mathcal{H}$

$$\sup_{n,m \geq N} \|f_n - f_m\|_{\mathcal{H}} \xrightarrow[N \to +\infty]{} 0 \Rightarrow \sup_{n,m \geq N} \|\mathcal{V}_0^{-1} f_n - \mathcal{V}_0^{-1} f_m\|_{\mathcal{F}} \xrightarrow[N \to +\infty]{} 0$$

so $(\mathcal{V}_0^{-1} f_n)_{n \geq 0}$ is a Cauchy sequence in $\mathcal{F}$, so it converges in $\mathcal{F}$.

Therefore, there exists $w \in \mathcal{F}$, such that

and $w_0 \in \mathcal{F}_0$

$(\nu_0^{-1} f_n)$ converges towards $w = w_0 + (w - w_0)$ where $w - w_0 \in \mathcal{F}_0^{\perp}$
$\implies 0 \leq \|\nu_0^{-1} f_n - w\|_{\mathcal{F}} \to 0$
$\implies 0 \leq \|\nu_0^{-1}(f_n - \nu_0 w_0) - (w - w_0)\|_{\mathcal{F}} \to 0$ because $w_0 \in \mathcal{F}_0$
$\implies 0 \leq \|f_n - \nu_0 w_0\|_{\mathcal{H}} + \|(w - w_0)\|_{\mathcal{F}} \to 0$ because of Pythagora's theorem
$\implies \|f_n - \nu_0 w_0\|_{\mathcal{H}} \to 0$ (and $w = w_0$)
Therefore (f_n) converges towards $\nu_0 w_0 \in \mathcal{H}$

So $\mathcal{H}$ is complete for $\|.\|_{\mathcal{H}}$
That concludes that $\mathcal{H}$ is a Hilbert space

- Let $x \in \mathcal{X}$, and $K_x : t \mapsto K(x,t) = \langle \varphi(x), \varphi(t) \rangle_{\mathcal{F}}$
  $\forall t \in \mathcal{X}$, $K_x(t) = f_{\varphi(x)}(t) \Rightarrow K_x = f_{\varphi(x)}$ so $K_x \in \mathcal{H}$

- Let $x \in \mathcal{X}$ and $f \in \mathcal{H}$. There exists $w_0 \in \mathcal{F}_0$ such that $f = f_{w_0}$
  $f_{w_0}(x) = \langle \varphi(x), w_0 \rangle_{\mathcal{F}}$
  Let $v \in \mathrm{Ker}\,\nu$, $\nu(v) = 0 \Rightarrow \forall x' \in \mathcal{X}, \langle \varphi(x'), v \rangle = 0$
  so $\langle \varphi(x), v \rangle = 0$
  $\Rightarrow \varphi(x) \in (\mathrm{Ker}\,\nu)^{\perp} = \mathcal{F}_0$
  so $f_{w_0}(x) = \langle \nu_0^{-1}\nu_0(\varphi(x)), \nu_0^{-1}\nu_0(w_0) \rangle = \langle \nu_0^{-1} K_x, \nu_0^{-1} f_{w_0} \rangle$
  $= \langle K_x, f \rangle_{\mathcal{H}}$
  so the reproducing property holds.

Therefore, $\mathcal{H}$ is the RKHS of K endowed with the norm: $\|f\|_{\mathcal{H}} = \|\nu_0^{-1} f\|_{\mathcal{F}}$

Exercise 2

Let K a p.d kernel on a space $\mathcal{X}$, and $f : \mathcal{X} \to \mathbb{R}$.

($\Rightarrow$) Assume f belongs to the RKHS $\mathcal{H}$ with kernel K.

- Let $N \in \mathbb{N}$, $(x_1, \dots, x_N) \in \mathcal{X}^N$ and $(a_1, \dots, a_N) \in \mathbb{R}^N$

$$\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j \left(K(x_i, x_j) - \lambda f(x_i) f(x_j) \right)$$
$$= \sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j K(x_i, x_j) - \lambda \sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j f(x_i) f(x_j)$$
$$= \sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j \langle K_{x_i}, K_{x_j} \rangle_{\mathcal{H}} - \lambda \sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j \langle K_{x_i}, f \rangle_{\mathcal{H}} \langle K_{x_j}, f \rangle_{\mathcal{H}}$$
$$= \langle \sum_{i=1}^{N} a_i K_{x_i}, \sum_{i=1}^{N} a_i K_{x_i} \rangle_{\mathcal{H}} - \lambda \langle \sum_{i=1}^{N} a_i K_{x_i}, f \rangle^2$$

by Cauchy-Schwarz inequality

$$\geq \| \sum_{i=1}^{N} a_i K_{x_i} \|_{\mathcal{H}}^2 - \lambda \|f\|_{\mathcal{H}}^2 \| \sum_{i=1}^{N} a_i K_{x_i} \|_{\mathcal{H}}^2$$
$$\geq \| \sum_{i=1}^{N} a_i K_{x_i} \|_{\mathcal{H}}^2 \left(1 - \lambda \|f\|_{\mathcal{H}}^2 \right)$$

- If $f = 0$, for $\lambda = 1$, $K(x,x') - \lambda f(x) f(x') = \underbrace{K(x,x')}_{\text{p.d}}$
  so $K(x,x') - \lambda f(x) f(x')$ is p.d.
- Otherwise, for $\lambda = 1/\|f\|_{\mathcal{H}}^2 > 0$, the previous calculation gives that $\forall N \in \mathbb{N}$, $(x_1, \dots, x_N) \in \mathcal{X}^N$, $(a_1, \dots, a_N) \in \mathbb{R}^N$
$$\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j \left(K(x_i, x_j) - \lambda f(x_i) f(x_j) \right) \geq 0$$

Moreover, $\forall x, x' \in \mathcal{X}$, $K(x,x') - \lambda f(x) f(x') = K(x',x) - \lambda f(x') f(x)$

Hence, $K(x,x') - \lambda f(x) f(x')$ is p.d

($\Leftarrow$) Assume there exists $\lambda > 0$ such that $K(x,x') - \lambda f(x) f(x')$ is p.d., and let $\mathcal{H}$ be the RKHS of kernel K.

- $\forall x, x' \in \mathcal{X}$ $\quad f(x) f(x') = f(x') f(x)$

  $\forall N \in \mathbb{N}, (x_1, \dots, x_N) \in \mathcal{X}^N, (a_1, \dots, a_N) \in \mathbb{R}^N$

$$\sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j f(x_i) f(x_j) = \left(\sum_{i=1}^{N} a_i f(x_i) \right)^2 \geq 0$$

so $\underline{x, x' \mapsto f(x) f(x') \text{ is a p.d. kernel on } \mathcal{X}.}$

- Recall from exercise 5.1 that:

  For any p.d. kernels K_1 and K_2 on $\mathcal{X}$ and $\alpha, \beta \in \mathbb{R}^*_+$, $\alpha K_1 + \beta K_2$ is also p.d., with RKHS $H_1 + H_2$ endowed with the norm:

$$\| f \|^2 = \min_{\substack{f_1 \in H_1,\, f_2 \in H_2 \\ f = f_1 + f_2}} \left\{ \frac{1}{\alpha} \| f_1 \|^2_{H_1} + \frac{1}{\beta} \| f_2 \|^2_{H_2} \right\}$$

Therefore, $K(x, x') = \underbrace{K(x,x') - \lambda f(x) f(x')}_{\text{p.d. kernel}} + \overbrace{\lambda}^{\in \mathbb{R}^+} \underbrace{f(x) f(x')}_{\text{p.d. kernel}}$

is a p.d. kernel with RKHS $\underline{H_1 + H_2}$ where H_1 is the RKHS of $K(x, x') - \lambda f(x) f(x')$ and H_2 the one of $f(x) f(x')$

- If $f = 0$, $f \in \mathcal{H}$ because $\mathcal{H}$ is a Hilbert space.

  Otherwise, $\exists x \in \mathcal{X}$ $f(x) \neq 0$.

  We know that H_2 contains $t \mapsto f(x) f(t)$

  so H_2 contains $1/f(x) \times (t \mapsto f(x) f(t))$, so H_2 contains f

  so $\underline{f \in H_1 + H_2}$.

- Finally, because of the uniqueness of the RKHS, $H_1 + H_2 = \mathcal{H}$.

  so $\underline{f \in \mathcal{H}}$.