Proposition 1. *Exponential kernel (also called Laplace kernel)* $k(\mathbf{x},\mathbf{y}) = e^{-\gamma\|\mathbf{x}-\mathbf{y}\|}$*, where* $\gamma > 0$ *and* $\mathbf{x},\mathbf{y} \in \mathbb{R}^d$ *is a kernel.*

Proof. The idea is to represent exponential kernel as a non-negative linear combination of Gaussian RBF kernels. To do this formally, we resort to inverse Laplace transform:

Lemma 2. *The inverse Laplace transform of* $e^{-\gamma\sqrt{s}}$ *is*

$$e^{-\gamma\sqrt{s}} = \int_0^{+\infty} \frac{\gamma e^{-\frac{\gamma^2}{4t}}}{2\sqrt{\pi}t^{3/2}} e^{-ts} dt. \tag{1}$$

This can be computed directly using Bromwich integral, which can be found in most complex analysis textbooks[1].

Because the integrand is non-negative when $t \in [0, +\infty)$, $k(\mathbf{x},\mathbf{y})$ can be expressed as a non-negative linear combination of Gaussian RBF kernels:

$$k(\mathbf{x},\mathbf{y}) = \frac{\gamma}{2\sqrt{\pi}} \int_0^{+\infty} \frac{e^{-\frac{\gamma^2}{4t}}}{t^{3/2}} e^{-t\|\mathbf{x}-\mathbf{y}\|^2} dt. \tag{2}$$

Due to the closure property of kernels, exponential kernel is a kernel.

Strictly speaking, the "linear combination" is an integral. To be mathematically rigorous, one can further Fourier transform the Gaussian RBF kernel, as it only depends on $\mathbf{x}-\mathbf{y}$:

Lemma 3. *The Fourier transform of a Gaussian is a Gaussian:*

$$e^{-\alpha\|\mathbf{x}\|^2} = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \frac{e^{-\frac{\|\mathbf{y}\|^2}{4\alpha}}}{(2\alpha)^{d/2}} e^{-i\mathbf{x}\cdot\mathbf{y}} d\mathbf{y}. \tag{3}$$

It follows:

$$\begin{aligned}
\sum_{i,j} c_i c_j k(\mathbf{x}_i,\mathbf{x}_j) =& \sum_{i,j} c_i c_j e^{-\gamma\sqrt{\|\mathbf{x}_i-\mathbf{x}_j\|^2}} & (4)\\
=& \frac{\gamma}{2\sqrt{\pi}} \int_0^{+\infty} dt \sum_{i,j} c_i c_j \frac{e^{-\frac{\gamma^2}{4t}}}{t^{3/2}} e^{-t\|\mathbf{x}_i-\mathbf{x}_j\|^2} dt & (5)\\
=& \frac{\gamma}{2\sqrt{\pi}} \int_0^{+\infty} dt \sum_{i,j} c_i c_j \frac{e^{-\frac{\gamma^2}{4t}}}{t^{3/2}} \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \frac{e^{-\frac{\|\mathbf{y}\|^2}{4t}}}{(2t)^{d/2}} e^{-i(\mathbf{x}_i-\mathbf{x}_j)\cdot\mathbf{y}} d\mathbf{y} & (6)\\
=& \frac{\gamma}{2^{d+1}\pi^{(d+1)/2}} \int_0^{+\infty} dt \int_{\mathbb{R}^d} d\mathbf{y} \frac{e^{-\frac{\gamma^2}{4t}}}{t^{(d+3)/2}} \left|\sum_i c_i e^{-i\mathbf{x}_i\cdot\mathbf{y}}\right|^2 \geq 0. & (7)
\end{aligned}$$

Therefore the exponential kernel is positive semi-definite and is indeed a kernel. □

The key in the above proof is that the Fourier transform of the exponential kernel is non-negative. To obtain the Fourier transform of the exponential kernel, we first compute its inverse Laplace transform and represent the result as an integral over t. In this way, our proof can be easily generalized to the following proposition

Proposition 4. *If a function* $\varphi(s)$ *has a non-negative inverse Laplace transform, i.e.*

$$\varphi(s) = \int_0^{+\infty} \mu(t) e^{-ts} dt, \tag{8}$$

$\mu(t) \geq 0$ *for* $t \in [0, +\infty)$*, then* $k(\mathbf{x},\mathbf{y}) = \varphi(\|\mathbf{x}-\mathbf{y}\|^2)$*, where* $\mathbf{x},\mathbf{y} \in \mathbb{R}^d$ *is a kernel.*

[1] See also, for example [this Mathematics Stack Exchange page]

For example, because the inverse Laplace transform of $\varphi(s) = (1+s)^{-1}$ is $\mu(t) = e^{-t}$, using Proposition 4,

$$k(\mathbf{x}, \mathbf{y}) = \frac{1}{1 + \|\mathbf{x} - \mathbf{y}\|^2}, \tag{9}$$

is a kernel. Similarly, the inverse Laplace transform of $\varphi(s) = (1+\sqrt{s})^{-1}$ is also non-negative, because

$$\frac{1}{1+\sqrt{s}} = \int_0^{+\infty} e^{-(1+\sqrt{s})t} dt = \int_0^{+\infty} dt e^{-t} \int_0^{+\infty} \frac{te^{-\frac{t^2}{4x}}}{2\sqrt{\pi}x^{3/2}} e^{-xs} dx \tag{10}$$

$$= \int_0^{+\infty} \underbrace{\left(\int_0^{+\infty} dt \frac{e^{-t} te^{-\frac{t^2}{4x}}}{2\sqrt{\pi}x^{3/2}} \right)}_{\geq 0} e^{-xs} dx, \tag{11}$$

where we have used Lemma 2. Therefore,

$$k(\mathbf{x}, \mathbf{y}) = \frac{1}{1 + \|\mathbf{x} - \mathbf{y}\|}, \tag{12}$$

is also a kernel.

Before concluding, we note that the inverse of Proposition 4 is also true, although the proof is much more non-trivial. We have the following characterizations of positive definite functions (note that all three theorems below are stated with "if and only if"), which were proved in the 1930s:

Definition 1 (Positive Definite Function)**.** A function $\Phi(\mathbf{x}) : \mathbb{R}^d \to \mathbb{C}$ is positive definite if for all $n \in \mathbb{N}$ and $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$, the matrix M, where $M_{ij} = \Phi(\mathbf{x}_i - \mathbf{x}_j)$, is positive semi-definite.

Theorem 5 (Bochner)**.** *A function $\Phi(\mathbf{x}) : \mathbb{R}^d \to \mathbb{C}$ is positive definite if and only if it has non-negative Fourier transform on $\mathbb{R}^d$.*

The above theorem characterizes all translation invariant positive definite functions. For radial functions, i.e. $\Phi(\mathbf{x}) = \varphi(\|\mathbf{x}\|)$, we have further characterizations:

Definition 2 (Complete Monotone Function)**.** A function on $\varphi : [0, +\infty) \to \mathbb{R}$ is complete monotone on $[0, +\infty)$ if it is continuous on $[0, +\infty)$, infinitely differentiable on $(0, +\infty)$, and satisfies

$$(-1)^n \frac{d^n}{dr^n} \varphi(r) \geq 0, \tag{13}$$

for all $n \in \mathbb{N}$ and $r > 0$.

Theorem 6 (Schoenberg)**.** *A function $\varphi(r)$ is complete monotone on $[0, +\infty)$ if and only if $\Phi = \varphi(\|\mathbf{x}\|^2)$ is positive definite on $\mathbb{R}^d$ for all d.*

Theorem 7 (Hausdorff-Bernstein-Widder)**.** *A function $\varphi : [0, +\infty) \to \mathbb{R}$ is complete monotone on $[0, +\infty)$ if and only if it has non-negative Laplace transform on $[0, +\infty)$.*