

# INSA – Gaussian processes

## Spectral representation and Bochner's theorem

---

Andrés F. López-Lopera

The French Aerospace Lab ONERA, France  
Information Processing and Systems Department (DTIS)  
Multidisciplinary Methods, Integrated Concepts (M2CI) Research Unit

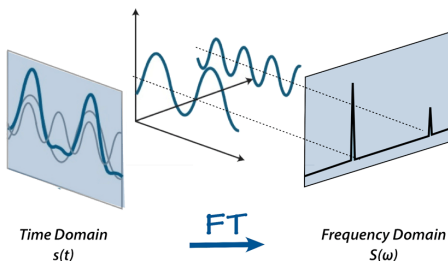
The background of the slide is a light gray with a series of overlapping, wavy lines in various colors including purple, blue, green, yellow, and orange. These lines create a sense of motion and depth. The text 'Spectral analysis' is centered on the left side of the slide, with a thin horizontal line extending to the right from its base.

## **Spectral analysis**

---

# Spectral analysis

- Spectral methods are widely used for data analysis
- Analysis in terms of a spectrum of frequencies, energies, eigenvalues, etc

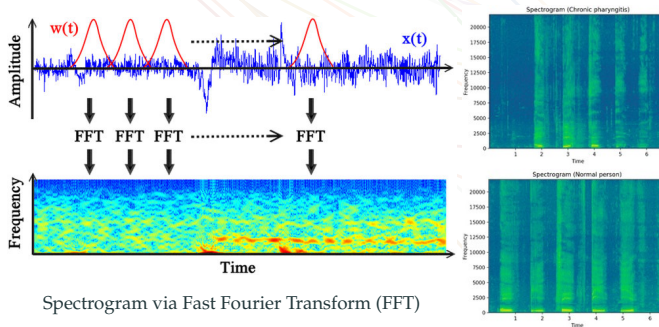


Fourier analysis

# Spectral analysis

- Spectral analysis has its root in communications but it is also used in:
  - Electrical engineering
  - Acoustical engineering
  - Materials science
  - Geophysics
  - Atmospheric science & astronomy
  - ...

Speech spectrum of patients with chronic pharyngitis



Spectrogram via Fast Fourier Transform (FFT)

- Spectral analysis is a powerful tool to study complex-valued functions:

$$z(x) = u(x) + iv(x),$$

with  $x \in \mathbb{R}^d$ ,  $i = \sqrt{-1}$  (imaginary number), and  $u, v$  real-valued functions.

- Spectral analysis is a powerful tool to study complex-valued functions:

$$z(x) = u(x) + iv(x),$$

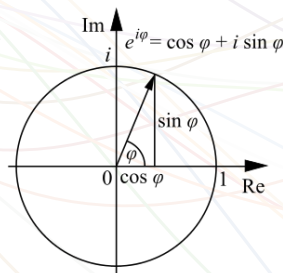
with  $x \in \mathbb{R}^d$ ,  $i = \sqrt{-1}$  (imaginary number), and  $u, v$  real-valued functions.

- An example of a complex-valued function is the **Euler's formula**:

$$z(\varphi) = \cos(\varphi) + i \sin(\varphi) = e^{i\varphi}, \quad \text{for } \varphi \in \mathbb{C}.$$

**Euler's identity:**

$$e^{i\varphi} - 1 = 0.$$



- In Fourier analysis, the spectral (frequency) representation of  $z : \mathbb{R}^d \rightarrow \mathbb{C}$  is given by the **Fourier transform (FT)**:

$$S(\omega) := \mathcal{F}[f](\omega) = \int_{\mathbb{R}^d} \exp(-i\langle \omega, x \rangle) z(x) dx,$$

where  $\langle \cdot, \cdot \rangle$  is the scalar product on  $\mathbb{R}^d$ .

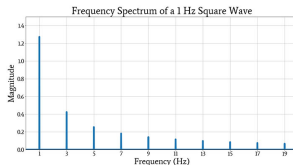
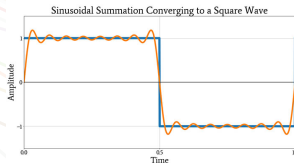
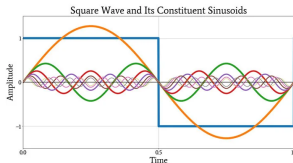
- The **inverse Fourier transform** of  $S(\omega)$  is given by

$$\mathcal{F}^{-1}[S](x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp(i\langle \omega, x \rangle) S(\omega) d\omega.$$

## Fourier series (sine-cosine form):

$$z_{n_{\text{freq}}}(x) = \sum_{n=-n_{\text{freq}}}^{n_{\text{freq}}} c_n \exp(i\omega_n x) = \sum_{n=-n_{\text{freq}}}^{n_{\text{freq}}} c_n [\cos(\omega_n x) + i \sin(\omega_n x)]$$

with  $x \in \mathbb{R}$  and  $\omega_n = n\omega_0 = 2\pi n f_0$



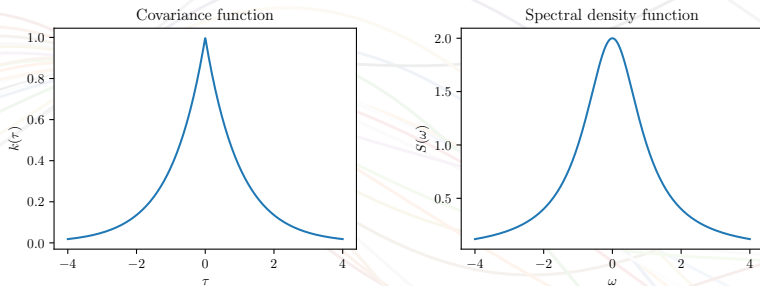
Fourier representation of the square wave function

[square wave] [sawtooth wave]



# Spectral representation of kernels

- Fourier analysis can also be applied to random fields and kernels.



Spectral representation of the exponential kernel

- The spectral representation can be used for kernel design [Heinonen, 2017]:
  - All stationary kernels  $k(\tau)$  have a spectral density  $S(\omega)$
  - All spectral densities  $S(\omega)$  define a covariance function

1. Complex-valued random fields
2. Spectral representation of random fields
3. Bochner's theorem
4. Spectral kernel design

- In the following, we consider:
  - Stationary (centred) random fields  $\{Y(\mathbf{x}); \mathbf{x} \in \mathbb{R}^d\}$  (e.g. a GP)
  - Stationary kernels  $k(\boldsymbol{\tau}) := k(\mathbf{x}, \mathbf{x} + \boldsymbol{\tau})$  (abuse of notation)

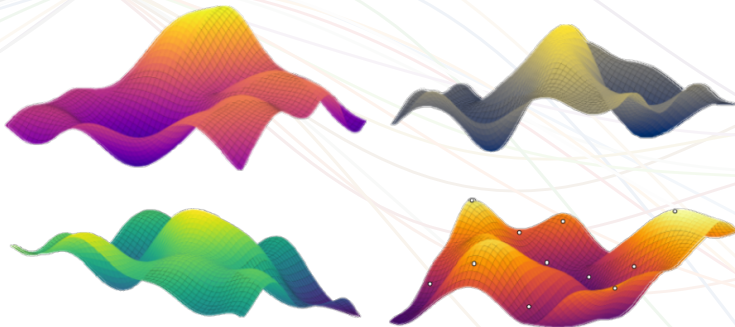
The background of the slide features a series of smooth, flowing, wavy lines in various colors including light blue, green, yellow, orange, and purple. These lines originate from the left side and curve across the frame, creating a sense of movement and depth. The lines are thin and have a soft, ethereal quality.

## **Complex-valued random fields**

---

- A random field is a random function over an arbitrary domain, e.g. in  $\mathbb{R}^d$ :

Gaussian random field:  $Y(x) \sim \mathcal{GP}(0, k)$ .



2D Gaussian random fields

- A random field  $\{Z(x); x \in \mathbb{R}^d\}$  is complex-valued if:

$$Z(x) = U(x) + iV(x),$$

where  $i = \sqrt{-1}$  and  $U, V$  are real-valued random fields.

- We define  $\bar{Z}$  as the complex conjugate of  $Z$ :

$$\overline{Z(x)} = \overline{U(x) + iV(x)} = U(x) - iV(x)$$

- We denote  $\text{Re}\{\cdot\}$  and  $\text{Im}\{\cdot\}$  operators such that:

$$\text{Re}\{Z(x)\} = U(x), \quad \text{Im}\{Z(x)\} = V(x).$$

- The expectation of a complex-valued random field  $Z$  is given by:

$$\mathbb{E} \{Z(x)\} = \mathbb{E} \{U(x) + iV(x)\} = \mathbb{E} \{U(x)\} + i\mathbb{E} \{V(x)\}.$$

- Note that  $\mathbb{E} \{Z(x)\}$  exists iff  $\mathbb{E} \{U(x)\}$  and  $\mathbb{E} \{V(x)\}$  exist.
- If  $\mathbb{E} \{Z(x)\}$  exists, then the complex conjugation of  $\mathbb{E} \{Z(x)\}$  is given by

$$\overline{\mathbb{E} \{Z(x)\}} = \mathbb{E} \left\{ \overline{Z(x)} \right\} = \mathbb{E} \{U(x)\} - i\mathbb{E} \{V(x)\}.$$

- The covariance of a operation of complex-valued random field  $Z$  is given by:

$$\text{cov} \left\{ Z(x), \overline{Z(x')} \right\} = \mathbb{E} \left\{ Z(x) \overline{Z(x')} \right\}$$

**Exercise.** Compute  $k(\tau) = \text{cov} \left\{ Z(x + \tau), \overline{Z(x)} \right\}$ .

**Exercise.** Show that  $k(-\tau) = \overline{k(\tau)}$ .



## Definition (positive semidefiniteness for complex-valued functions)

A kernel  $k : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{C}$  is a p.s.d. complex function if for all  $n \in \mathbb{N}$ , and for all  $c_1, \dots, c_n \in \mathbb{C}, \forall \mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$ ,

$$\sum_{p=1}^n \sum_{q=1}^n c_p \bar{c}_q k(\mathbf{x}_p - \mathbf{x}_q) \geq 0.$$

## Definition (positive semidefiniteness for real-valued functions)

A kernel  $k : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  is a p.s.d. function if for all  $n \in \mathbb{N}$ , and for all  $a_1, \dots, a_n \in \mathbb{R}, \forall \mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$ ,

$$\sum_{p=1}^n \sum_{q=1}^n a_p a_q k(\mathbf{x}_p - \mathbf{x}_q) \geq 0.$$

The background of the slide features a series of overlapping, wavy lines in various colors including light blue, green, yellow, orange, and red. These lines flow from the left side towards the right, creating a sense of movement and depth. The lines are thin and have a soft, ethereal quality.

## **Spectral representation of random fields**

---

- The spectral representation requires the following conditions:

## Definition (Mean square continuity – see, e.g., [Stein, 1999])

A complex-valued random field  $Z$  is mean square continuous at  $x$  if:

$$\lim_{x' \rightarrow x} \mathbb{E} \left\{ |Z(x) - Z(x')|^2 \right\} = 0.$$

## Definition (Weak stationarity – see, e.g., [Stein, 1999])

A complex-valued random field  $\{Z(x); x \in \mathbb{R}^d\}$  is called *weakly stationary* if:

- $\mathbb{E} \{Z(x)\}$  is constant
- $\mathbb{E}\{Z(x)\overline{Z(x)}\} < \infty$  (finite second-order moments)
- $\text{cov}\{Z(x), \overline{Z(x + \tau)}\} = k(\tau)$  (stationary kernel)

**Note:** a strictly stationary random field with finite second-order moments is also weakly stationary.

**Exercise.** Suppose  $\omega_1, \dots, \omega_n \in \mathbb{R}^d$  and let  $Z_1, \dots, Z_n$  be zero-mean complex random variables with  $\mathbb{E} \{Z_p \overline{Z_q}\} = 0$  for  $p \neq q$  and  $\mathbb{E} \{|Z_p|^2\} = f_p < \infty$ . Consider

$$\tilde{Z}(x) = \sum_{p=1}^n Z_p \exp(i\langle \omega_p, x \rangle).$$

Show that  $\tilde{Z}$  is weakly stationary.

- $\tilde{Z}(x) = \sum_{p=1}^n Z_p \exp(i\langle \omega_p, x \rangle)$  is an example of a spectral representation.
- Generally, spectral representations of mean square continuous weakly stationary random fields can be obtained [Stein, 1999]:

$$Z(x) = \int_{\mathbb{R}^d} \exp(i\langle \omega, x \rangle) M(d\omega) \quad (1)$$

where  $M$  is a *complex random measure* on  $\mathbb{R}^d$ .

- The integral in (1) can be seen as a limit in  $L^2$  of the sums used for  $\tilde{Z}$ .

**Exercise.** Suppose that for some positive finite measure  $\mu$ :

$$\mathbb{E} \{M(\Delta)\} = 0$$

$$\mathbb{E} \left\{ |M(\Delta)|^2 \right\} = \mu(\Delta)$$

$$\mathbb{E} \left\{ M(\Delta_1), \overline{M(\Delta_2)} \right\} = 0, \quad \text{for disjoint Borel sets } \Delta_1, \Delta_2$$

Compute  $k(\tau) = \text{cov} \left\{ Z(x), \overline{Z(x + \tau)} \right\}$ .

**Solution.**

$$\begin{aligned} k(\tau) &= \text{cov} \left\{ \int_{\mathbb{R}^d} \exp(i\langle \omega, x \rangle) M(d\omega), \overline{\int_{\mathbb{R}^d} \exp(i\langle \nu, x + \tau \rangle) M(d\nu)} \right\} \\ &= \int_{\mathbb{R}^d} \exp(i\langle \omega, x \rangle) \int_{\mathbb{R}^d} \exp(i\langle \nu, x + \tau \rangle) \text{cov} \left\{ M(d\omega), \overline{M(d\nu)} \right\} \\ &= \int_{\mathbb{R}^d} \exp(i\langle \omega, \tau \rangle) \mu(d\omega) \end{aligned}$$

□

The background of the slide is a light gray with a series of overlapping, wavy lines in various colors including purple, blue, green, yellow, and orange. These lines create a sense of movement and depth. The text 'Bochner's theorem' is centered on the left side of the slide, with a horizontal orange line extending to the right from its base.

## Bochner's theorem

---

## Theorem (Bochner's theorem – see, e.g., [Stein, 1999])

· A complex-valued function  $k$  on  $\mathbb{R}^d$  is the covariance function of a weakly stationary mean square continuous complex-valued random process on  $\mathbb{R}^d$  iff it can be represented as

$$k(\tau) = \int_{\mathbb{R}^d} \exp(i\langle \omega, \tau \rangle) \mu(d\omega),$$

where the positive finite measure  $\mu$  is known as the *spectral measure*.

**Proof.** See previous exercise.



· **Exercise.** Let  $k$  be a real function on  $\mathbb{R}^d$ . Verify that the kernels of the form

$$k(x, x') = \int_{\mathbb{R}^d} \cos(2\pi \langle s, x - x' \rangle) \mu(ds),$$

are positive semidefinite (p.s.d.).

## Theorem (Wiener-Khintchine theorem – see, e.g., [Chatfield, 2016])

· If  $\mu$  has a density  $S$  (known as the spectral density of  $k$ ), then:

$$k(\tau) = \int_{\mathbb{R}^d} \exp(i\langle \omega, \tau \rangle) S(\omega) d\omega.$$

· If  $S$  exists, we have the inversion formula:

$$S(\omega) = \int_{\mathbb{R}^d} \exp(-i\langle \omega, \tau \rangle) k(\tau) d\tau.$$

**Proof.** Apply Bochner's theorem and the Fourier's formulas.

**Example.** Consider the spectral density:

$$S(\omega) = \delta(\omega - \omega_0),$$

with  $\omega \in \mathbb{R}$ . Compute the corresponding valid kernel  $k(\tau)$ .

**Solution.**

$$k(\tau) = \int_{\mathbb{R}} \exp(i\omega\tau) \delta(\omega - \omega_0) d\omega = \exp(i\omega_0\tau) = \cos(\omega_0\tau) + i \sin(\omega_0\tau).$$

Note that, if considering a real-valued random field, then

$$k(\tau) = \cos(\omega_0\tau).$$

□

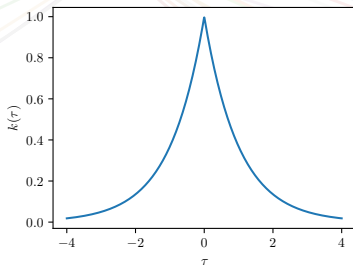
**Exercise.** Compute the spectral density of the kernel:

$$k(\tau) = \exp\left(-\frac{|\tau|}{\ell}\right),$$

with  $\tau \in \mathbb{R}$  and  $\ell > 0$ .

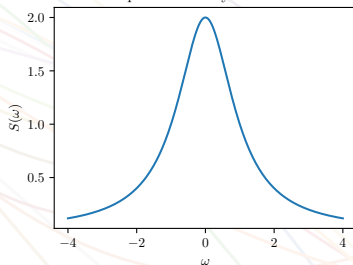
$$k(\tau) = \exp\left(-\frac{|\tau|}{\ell}\right)$$

Covariance function

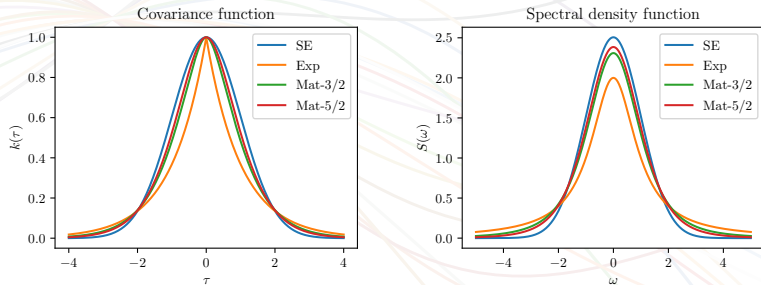


$$S(\omega) = \frac{2}{\frac{1}{\ell} + \ell\omega^2}$$

Spectral density function



Spectral representation of the exponential kernel



Examples of spectral density functions [Heinonen, 2017]

- Bochner's theorem can be used to prove p.s.d. of usual stationary kernels
  - The SE is the FT of a Gaussian function, i.e. it is p.s.d.
  - Matérn kernels are FT of Student  $t$  functions, i.e. they are p.s.d.
- It can also be generalised to distributions:
  - $\delta_{x-x'}$  is the FT of the constant functions, i.e. it is p.s.d.
  - The constant function is the FT of  $\delta_{x-x'}$ , i.e. it is p.s.d.

# 1D stationary kernel functions

- Some classic kernels for stationary processes:

Stationary kernel name	$k(\tau) := k(x, x + \tau)$	Spectral density
Cosine	$\sigma^2 \cos(2\pi\tau)$	Dirac delta
Sinc	$\sigma^2 \frac{\sin(\pi\tau)}{\tau}$	Uniform
Squared Exponential	$\sigma^2 \exp\left\{-\frac{1}{2} \frac{\tau^2}{\ell^2}\right\}$	Gaussian
Exponential	$\sigma^2 \exp\left\{-\frac{ \tau }{\ell}\right\}$	Student $t_{1/2}$
Matérn 3/2	$\sigma^2 \left(1 + \sqrt{3} \frac{ \tau }{\ell}\right) \exp\left\{-\sqrt{3} \frac{ \tau }{\ell}\right\}$	Student $t_{3/2}$
Matérn 5/2	$\sigma^2 \left(1 + \sqrt{5} \frac{ \tau }{\ell} + \frac{5}{3} \frac{\tau^2}{\ell^2}\right) \exp\left\{-\sqrt{5} \frac{ \tau }{\ell}\right\}$	Student $t_{5/2}$

**Exercise.** Compute the spectral density  $S(\omega)$  for the previous kernels.



## Extension to non-stationary kernels

- Non-stationary assumptions imply:  $k(x, x') \neq k(x - x')$
- Examples of non-stationary kernels are [Heinonen et al., 2016]

$$k(x, x') = \langle x, x' \rangle^d \quad (\text{polynomial kernel})$$

$$k(x, x') = w(x)w(x') \sqrt{\frac{2\ell(x)\ell(x')}{\ell^2(x) + \ell^2(x')}} \exp\left(\frac{\|x - x'\|^2}{\ell^2(x) + \ell^2(x')}\right) \quad (\text{Gibbs kernel})$$

### Definition (Bochner's theorem for non-stationary kernels [Heinonen, 2017])

- If  $\mu$  has a spectral density  $S$  of  $k$ ), then:

$$k(x, x') = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \exp(i[\langle \omega, \tau \rangle - \langle \omega', \tau' \rangle]) S(\omega, \omega') d\omega d\omega'.$$

- If  $S$  exists, we have the inversion formula:

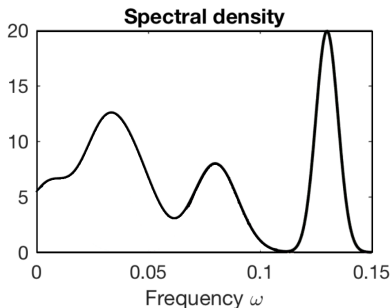
$$S(\omega, \omega') = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \exp(-i[\langle \omega, \tau \rangle - \langle \omega', \tau' \rangle]) k(x, x') dx dx'.$$

The background of the slide is a light gray with a series of overlapping, wavy lines in various colors including purple, blue, green, yellow, and orange. These lines create a sense of movement and depth. The text 'Spectral kernel design' is centered on the left side of the slide, with a horizontal orange line extending to the right from its base.

## Spectral kernel design

---

- Remember that all the spectral densities  $S(\omega)$  define a covariance function
- Then one can use  $S(\omega)$  for the design of stationary kernels  $k(\tau)$



Example of a spectral density [Heinonen, 2017]

## Sparse-spectrum kernel [Lázaro-Gredilla et al., 2010]:

- One can consider the spectral density given by

$$S(\omega) = \frac{1}{Q} \sum_{q=1}^Q \delta(\omega - \omega_q),$$

with real frequencies  $\omega_1, \dots, \omega_Q \in \mathbb{R}$

- By applying the Wiener-Khintchine theorem, and using the Fourier dual

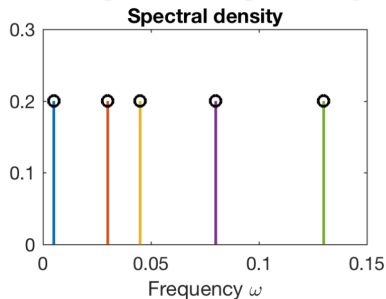
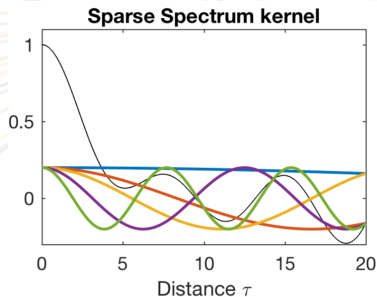
$$\tilde{S}(\omega) = \delta(\omega - \omega_0), \quad \tilde{k}(\tau) = \cos(\omega_0 \tau),$$

we have that

$$k(\tau) = \frac{1}{Q} \sum_{q=1}^Q \cos(\omega_q \tau),$$

with  $\tau \in \mathbb{R}$ .

## Sparse-spectrum kernel [Lázaro-Gredilla et al., 2010]



Spectral mixture kernel [Heinonen, 2017]

**Note.** Highly structured covariance (prone to overfitting) [Heinonen, 2017].

## Spectral mixture kernel [Wilson and Adams, 2013]:

- Define the real frequencies  $\omega_1, \dots, \omega_Q \in \mathbb{R}$  as Gaussian r.v.'s:

$$\omega_q \sim \mathcal{N}(m_q, \sigma_q^2), \quad \text{for } q = 1, \dots, Q,$$

with mean  $m_q$  and variance  $\sigma_q^2$ .

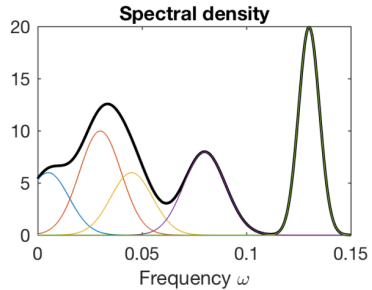
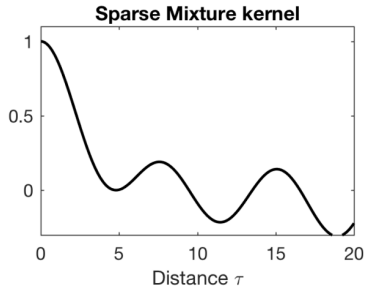
- Then, one can define the spectral density as a mixture of Gaussians:

$$S(\omega) = \sum_{q=1}^Q \alpha_q \mathcal{N}(m_q, \sigma_q^2).$$

- By applying the Wiener-Khintchine theorem, we have that

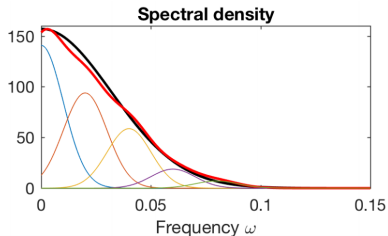
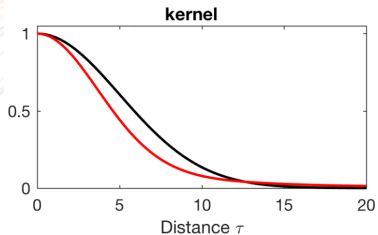
$$k(\tau) = \sum_{q=1}^Q \alpha_q \underbrace{\exp(-\sigma_q^2 \tau^2)}_{\text{smooth decay}} \underbrace{\cos(m_q \tau)}_{\text{periodic}},$$

with  $\tau \in \mathbb{R}$ .



Spectral mixture kernel [Heinonen, 2017]

# Spectral mixture kernel [Wilson and Adams, 2013]



■ SE kernel vs ■ Spectral Mixture kernel with  $Q = 5$  [Heinonen, 2017]

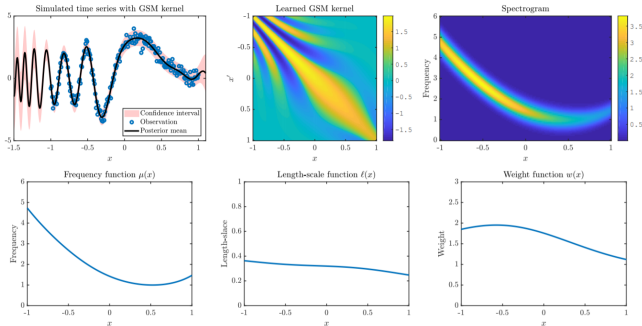


- A non-stationary version of the spectral mixture kernel is given by

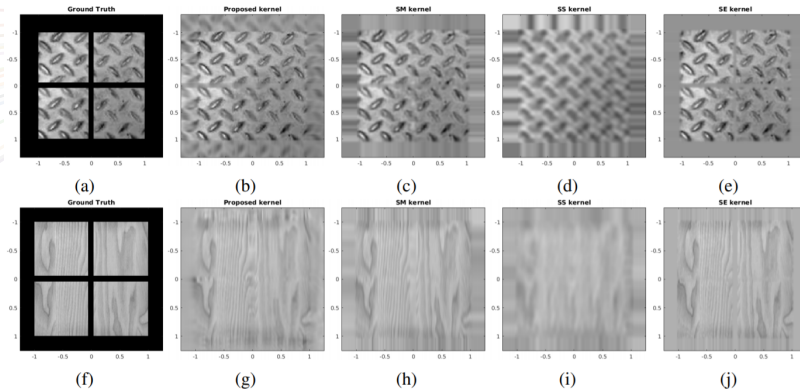
$$k(x, x') = \sum_{q=1}^Q \underbrace{w_q(x)w_q(x')}_{\text{smooth decay}} \underbrace{\exp(-[x\rho_q^2(x) + x'\rho_q^2(x')])}_{\text{smooth decay}} \underbrace{\langle \psi_q(x), \psi_q(x') \rangle}_{\text{periodic}},$$

with  $\psi_q(x) = [\cos(x\nu_q(x)), \sin(x\nu_q(x))]^\top$ , and

$$\log w_q(x) \sim \mathcal{GP}(0, k_w), \quad \log \rho_q(x) \sim \mathcal{GP}(0, k_\rho) \quad \log \nu_q(x) \sim \mathcal{GP}(0, k_\nu)$$



# Generalised spectral mixture kernel [Heinonen et al., 2016]



texture image examples [Heinonen et al., 2016]

The background of the slide is a light gray with a series of overlapping, wavy lines in various colors including purple, blue, green, yellow, and orange. These lines create a sense of movement and depth. The word "Conclusions" is centered on the left side of the slide, with a horizontal orange line extending to the right from its base.

## Conclusions

---

- Spectral analysis is a powerful tool to study complex-valued random fields

$$Z(x) = U(x) + iV(x).$$

- An example of a spectral representation is the Fourier transform
- All stationary kernels  $k(\tau)$  have a spectral density  $S(\omega)$

$$S(\omega) = \int_{\mathbb{R}^d} \exp(-i\langle \omega, \tau \rangle) k(\tau) d\tau.$$

- All spectral densities  $S(\omega)$  define a covariance function (Bochner's theorem)

$$k(\tau) = \int_{\mathbb{R}^d} \exp(i\langle \omega, \tau \rangle) S(\omega) d\omega.$$

- Spectral densities can be used for kernel design

- Jorge Buescu and A. C. Paixão. Real and complex variable positive definite functions. *São Paulo Journal of Mathematical Sciences*, 2012.
- C. Chatfield. *The Analysis of Time Series: An Introduction, Sixth Edition*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, 2016.
- Marc G. Genton. Classes of kernels for machine learning: A statistics perspective. *Journal of Machine Learning Research*, 2001.
- Markus Heinonen. Learning with spectral kernels. *Aalto University*, 2017.
- Markus Heinonen, Henrik Mannerström, Juho Rousu, Samuel Kaski, and Harri Lähdesmäki. Non-stationary Gaussian process regression with Hamiltonian Monte Carlo. *Proceedings of Machine Learning Research*, 2016.
- Miguel Lázaro-Gredilla, Joaquin Quiñonero-Candela, Carl E. Rasmussen, and Aníbal Figueiras-Vidal. Sparse spectrum Gaussian process regression. *Journal of Machine Learning Research*, 2010.
- Kevin P. Murphy. *Probabilistic Machine Learning: An introduction*. MIT Press, 2021.
- Carl E. Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. MIT Press, 2005.
- Arno Solin. Machine learning with signal processing. *ICML – TUTORIAL*, 2020.
- Michael L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, 1999.
- Andrew Gordon Wilson and Ryan Prescott Adams. Gaussian process kernels for pattern discovery and extrapolation. *Proceedings of Machine Learning Research*, 2013.