Nobel prize 2996 to John C. Mather and George F. Smoot “for discovery of the blackbody form and anisotropy of the cosmic microwave background radiation”

Ocean waves
The oceans cover 72% of the earth’s surface. Essential for life on earth, and huge economic importance through fishing, transportation, oil and gas extraction

fMRI brain scan

PET brain scan
Slides for Spectral Domain by F&R
Aliasing (F&R p. 60-65)

\[ f(t) \text{ a } (N, 1) \text{ field} \]

\[ \Delta = (\delta_1, \ldots, \delta_N) \in \mathbb{R}^N_+ \]

\[ z_1 = (z_{1,1}, \ldots, z_{1,N}), z_2 = (z_{2,1}, \ldots, z_{2,N}) \in \mathbb{Z}^N \]

Then, with coordinatewise operations,

\[
\exp\{ i(\lambda + z_2 2\pi / \Delta)(z_1 \Delta)^t \} \\
= \exp\{ i\lambda (z_1 \Delta)^t \} \exp\{ i z_2 z_1^t 2\pi \} = \exp\{ i\lambda (z_1 \Delta)^t \}
\]

so oscillations with frequency \( \lambda \) and frequency \( \lambda + z_2 2\pi / \Delta \) cannot be distinguished from one another from observations on the discrete grid \( \Delta \mathbb{Z}^N = \delta_1 \mathbb{Z} \times \cdots \times \delta_N \mathbb{Z} \). This is the alias effect.
Spectral distribution of sampled process (F&R p. 62)

If \( f \) has a spectral density \( g \), then the covariance function of the sampled process has the spectral representation

\[
C(\Delta t) = \int_{-\frac{2\pi}{\Delta t}}^{\frac{2\pi}{\Delta t}} e^{i\langle t, \lambda \rangle} g_{\Delta}(\lambda) d\lambda, \quad \text{for } t \in \mathbb{Z}^N
\]

with

\[
g_{\Delta}(\lambda) = \sum_{z \in \mathbb{Z}^2} g(\lambda + \frac{2\pi z}{\Delta})
\]
The alias effect for $N = d = 1$. The Nyquist frequency is $\pi/\Delta$. (The power in a frequency area $A$ is $\nu(A)$.) If possible the sampling grid $\Delta$ should be fine enough to make the power outside of the Nyquist frequency small.
These are spectral densities for $(N,1)$ fields of the form

$$f(\lambda) = \phi(\alpha^2 + \lambda^2)^{(-\nu-N/2)},$$

with parameters $\phi, \alpha, \nu > 0$. Different expressions for the corresponding covariance functions exist. Classical spatial statistics by and large consists of fitting a Gaussian field with a Matern covariance to data, and to use the fitted model for kriging (=spatial prediction)

$\nu$ determines the smoothness of the field; is difficult to estimate; and is the most influential parameter for kriging

$\alpha^{-1}$ determines the “correlation range”
**FIGURE 5.3**
Matérn covariance functions. The solid line represents a Matérn covariance with $\nu = 1/2$ (exponential covariance) and the dashed line a Matérn with $\nu = 3/2$.

**FIGURE 5.4**
Spectral densities. The solid line represents a Matérn spectral density with $\nu = 1/2$ (the corresponding covariance function is exponential), and the dashed line a Matérn spectral density with $\nu = 3/2$. 
An alternative: the powered exponential model

\[ C(h) = \sigma e^{-\alpha |h|^\nu} \]

**FIGURE 5.2**
Covariance models: Exponential (solid line) and squared exponential also called Gaussian covariance (dotted line). (Corresponding to \( \nu = 1 \) and \( \nu = 2 \))
The periodogram (F&R p. 68-70)

\[ f(t) \text{ a (2,1) field, } \Delta = (\delta_1, \delta_2) \in \mathbb{R}_+^2, \ n = n_2 \]

Assume \( f \) is observed on a regular grid \( D(n_1 \times n_2) \) with spacing \( \delta_1 \) and number of points \( n_1 \) in the first coordinate, and spacing \( \delta_2 \) and number of points \( n_2 \) in the second coordinate.

The \textit{discrete Fourier transform of} \( f \):

\[
J(\lambda) = (\delta_1 \delta_2)^{-1/2} (n_1 n_2)^{-1/2} (2\pi)^{-1} \sum_{t_1=1}^{n_1} \sum_{t_2=1}^{n_2} f(\Delta t^t) e^{-i \Delta t \lambda^t}
\]

The \textit{periodogram}:

\[
I_n(\lambda) = |J(\lambda)|^2
\]
Calculation shows that

\[ I_n(\lambda) = \delta_1 \delta_2 (2\pi_2)^{-2} \sum_{t_1=1}^{n_1} \sum_{t_2=1}^{n_2} c_n(\Delta t^t) e^{-i\Delta t \lambda^t} \]

with

\[ c_n(\Delta h^t) = n^{-1} \sum_{t_1=1}^{n_1} \sum_{t_2=1}^{n_2} f(\Delta t^t)f(\Delta (t + h)^t) 1_{\{t + h \in D(n_1 \times n_2)\}} \]

the sample covariance function (assuming a mean zero field)
Increasing domain asymptotics: \( n_1, n_2 \to \infty, \; \frac{n_1}{n_2} \to c > 0 \) + moment conditions implies that

\[
I_n(\lambda_j) \to_d g_\Delta(\lambda) \chi^2_2 / 2 \quad \text{for} \; \lambda_j = \Delta j, \; \text{with} \; j \in \mathbb{Z}^2
\]

\( I_n(\lambda_j), I_n(\lambda_k) \) are asymptotically independent if \( j \neq k \)

Thus the periodogram is asymptotically unbiased estimator of \( g_\Delta(\lambda) \) (but not of \( g(\lambda) \)), but it is not consistent (why?)

Asymptotics for multiple independent observations on the same domain can also be developed. Bias is then the main problem.
Inconsistency is traditionally handled by *smoothing*:

i.e. by using a weighted average of the periodogram values in some neighborhood of $\lambda_j$ to estimate $g_\Delta(\lambda_j)$. This leads to more bias, but instead to smaller variance. The more observations one has the smaller one can make the neighborhood, and in this way obtain consistent estimators, where the bias and the variance tend to 0.

(This is similar to probability density estimation, and in fact precedes it)
Even if the periodogram is asymptotically unbiased, it can be severely biased for finite samples. One way of explaining the bias comes is that the sample covariance function is a bad estimator of the covariance for large $h$. Another is the formula

$$E(I_n(\lambda_0)) = (n_1n_2)^{-1}(2\pi)^{-2}\int_{\Pi_\Delta^2} g_\Delta(\lambda)W_n(\lambda - \lambda_0)d\lambda$$

with $\Pi_\Delta^2 = \left[-\frac{\pi}{\delta_1}, \frac{\pi}{\delta_1}\right] \times \left[-\frac{\pi}{\delta_2}, \frac{\pi}{\delta_2}\right]$ and $W_n$ the Fejér kernel

$$W_n(\lambda) = \prod_{j=1,2} \frac{\sin^2\left(\frac{n_j\lambda_j}{2}\right)}{\sin^2\left(\frac{\lambda_j}{2}\right)}$$
The bias can be reduced by windowing or data tapering: to replace \( f(t) \) in the definition of the periodogram by

\[
h(t) f(t)
\]

for a suitable function \( h(t) \) which small when \( t \) is close to the boundaries of the observation domain \( D(n_1 \times n_2) \).
If a parametric model \( \{g(\lambda; \theta) : \theta \in \Theta\} \) is assumed then \( \theta \) can be estimated by fitting \( \{g(\lambda_j; \theta), \lambda_j \in D(n_1 \times n_2)\} \) to the periodogram values \( I_n(\lambda_j) \) using weighted least squares, with weights \( g(\lambda_j; \theta)^{-1} \) (\( \theta \) is unknown, so the weights it has to be replaced by an estimate, perhaps iteratively)

Approximate maximum likelihood fitting is also possible, see F&R