3. The Gaussian kernel

Of all things, man is the measure.
Protagoras the Sophist (480-411 B.C.)

3.1 The Gaussian kernel

The Gaussian (better Gaußian) kernel is named after Carl Friedrich Gauß (1777-1855), a brilliant German mathematician.

```
<< "FrontEndVision`FEV6";
<< MathVisionTools`;
Needs["BarCharts"];
Show[Import["Gauss10DM.gif"], ImageSize -> 280]
```

Figure 3.1 The Gaussian kernel is apparent on the old German banknote of DM 10,- where it is depicted next to its famous inventor when he was 55 years old.

See also: http://scienceworld.wolfram.com/biography/Gauss.html.
The Gaussian kernel is defined in 1-D, 2D and N-D respectively as

\[
G_{1D}(x; \sigma) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{x^2}{2\sigma^2}}, \quad G_{2D}(x,y; \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad G_{ND}(\mathbf{x}; \sigma) = \frac{1}{(\sqrt{2\pi} \sigma)^N} e^{-\frac{||\mathbf{x}||^2}{2\sigma^2}}
\]

The \( \sigma \) determines the width of the Gaussian kernel.

### 3.2 Cascade property, selfsimilarity

The \textit{shape} of the kernel remains the same, irrespective of the \( \sigma \). When we \textit{convolve} two Gaussian kernels we get a new wider Gaussian with a variance \( \sigma^2 \) which is the sum of the variances of the constituting Gaussians:

\[
g_{\text{new}}(\mathbf{x}; \sigma_1^2 + \sigma_2^2) = g_1(\mathbf{x}; \sigma_1^2) \otimes g_2(\mathbf{x}; \sigma_2^2).
\]

\( \sigma = .; \text{ Simplify} \left[ \int_{-\infty}^{\infty} \text{gauss}[\alpha, \sigma_1] \text{gauss}[\alpha - x, \sigma_2] \, d\alpha \right] \)

\[
\int_{-\infty}^{\infty} \text{gauss}[\alpha, \sigma_1] \text{gauss}[-x + \alpha, \sigma_2] \, d\alpha
\]

### 3.3 The scale parameter

In order to avoid the summing of squares, one often uses the following parametrization:

\( 2 \sigma^2 \rightarrow t \), so the Gaussian kernel gets a particularly short form. In \( N \) dimensions:

\[
G_{ND}(\mathbf{x}, t) = \frac{1}{(\pi t)^{N/2}} e^{-\frac{||\mathbf{x}||^2}{2t}}.
\]

It is this \( t \) that emerges in the diffusion equation

\[
\frac{\partial L}{\partial t} = \frac{\partial^2 L}{\partial x^2} + \frac{\partial^2 L}{\partial y^2} + \frac{\partial^2 L}{\partial z^2}.
\]

It is often referred to as 'scale' (like in: differentiation to scale, \( \frac{\partial L}{\partial t} \)), but a better name is \textit{variance}.

We introduce the \textit{dimensionless} spatial parameter, \( \tilde{x} = \frac{x}{\sigma \sqrt{2}} \).

The new coordinate \( \tilde{x} = \frac{x}{\sigma \sqrt{2}} \) is called the \textit{natural coordinate}.

It eliminates the scale factor \( \sigma \) from the spatial coordinates, i.e. it makes the Gaussian kernels similar, despite their different inner scales. We will encounter natural coordinates many times hereafter.
3.4 Relation to generalized functions

When we take the limit as the inner scale goes down to zero (remember that $\sigma$ can only take positive values for a physically realistic system), we get the mathematical delta function, or Dirac delta function, $\delta(x)$.

This function is everywhere zero except in $x = 0$, where it has infinite amplitude and zero width, its area is unity.

$$\lim_{\sigma \to 0} \left( \frac{1}{\sqrt{2\pi \sigma}} e^{-\frac{x^2}{2\sigma^2}} \right) = \delta(x).$$

$\delta(x)$ is called the *sampling function* in mathematics. It is assumed that $f(x)$ is continuous at $x = a$:

$$\int_{-\infty}^{\infty} \text{DiracDelta}[x - a] f(x) \, dx$$

$f[a]$.

The *sampling property of derivatives* of the Dirac delta function is shown below:

$$\int_{-\infty}^{\infty} D[\text{DiracDelta}[x], \{x, 2\}] f(x) \, dx$$

$f''[0]$.

The integral of the Gaussian kernel from $-\infty$ to $x$ is a well known function as well. It is the *error function*, or *cumulative* Gaussian function, and is defined as:

$$\sigma = \ldots; \text{erf}[x_, \sigma_] = \int_{0}^{x} \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{y^2}{2\sigma^2}\right) \, dy$$

$$\frac{1}{2} \text{Erf}\left[\frac{x}{\sqrt{2} \sigma}\right]$$

$$\sigma = 1.\ldots; \text{Plot}\left[\frac{1}{2} \text{Erf}\left[\frac{x}{\sigma \sqrt{2}}\right], \{x, -4, 4\}, \text{AspectRatio} \to 0.3\right],$$

$$\text{AxesLabel} \to \{"x", \"Erf[\sigma]\"\}, \text{ImageSize} \to 300\right]$$

Erf[x]

For inner scale $\sigma \downarrow 0$, we get in the limiting case the *Heavyside function* or *unitstep function*. 
3.5 Relation to binomial coefficients

The Gaussian function emerges in expansions of powers of polynomials:

\[
\text{Manipulate}[\text{Plot} \left[ \frac{1}{2} \text{Erf} \left( \frac{x}{\sigma \sqrt{2}} \right) \right], (x, -4, 4), \text{AspectRatio} \to 0.3],
\]

\[
\text{AxesLabel} \to \{"x", \"Erf[x]\"\}, \text{ImageSize} \to 270], \{(\sigma, .4), .1, 2)\]

The coefficients of this expansion are the binomial coefficients \( \binom{n}{m} \) (‘\( n \) over ‘\( m \)’):

\[
\text{Manipulate}[\text{ListPlot}[\text{Table}[\text{Binomial[nr, n]}, \{n, 1, 100\}], \text{PlotStyle} \to \{\text{PointSize}[0.015]\}, \text{AspectRatio} \to 0.3], \{(nr, 30), 5, 100, 1)\]

\[
x^{10} + 10 x^9 y + 45 x^8 y^2 + 120 x^7 y^3 + 210 x^6 y^4 + 252 x^5 y^5 + 210 x^4 y^6 + 120 x^3 y^7 + 45 x^2 y^8 + 10 x y^9 + y^{10}
\]
Figure 3.2 Binomial coefficients approximate a Gaussian distribution for increasing order.

And here in two dimensions:

```math
BarChart3D@Table[Binomial[30, n] Binomial[30, m],
{n, 1, 30}, {m, 1, 30}], ImageSize -> 180]
```

### 3.6 The Fourier transform of the Gaussian kernel

The basis functions of the Fourier transform $\mathcal{F}$ are the sinusoidal functions $e^{i\omega x}$.

\[
\text{ExpToTrig}[e^{i\omega x}]
\]

\[
\cos(x \omega) + i \sin(x \omega)
\]

The definitions for the Fourier transform and its inverse are:

- the Fourier transform:
  \[
  F(\omega) = \mathcal{F}\{f(x)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{i\omega x} \, dx
  \]

- the inverse Fourier transform:
  \[
  F^{-1}\{F(\omega)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{-i\omega x} \, d\omega
  \]

\[
\sigma = \_, \text{gauss}[\omega_\_, \sigma_] =
\]

\[
\text{Simplify}\left[\frac{1}{\sqrt{2\pi}} \text{Integrate}\left[\frac{1}{\sigma \sqrt{2\pi}} \text{Exp}\left[-\frac{x^2}{2\sigma^2}\right] \text{Exp}[I \omega x], \{x, -\infty, \infty\}\right],
\{\sigma > 0, \text{Im}[\sigma] = 0\}\right]
\]

\[
\frac{e^{-\frac{1}{2} \omega^2 \sigma^2}}{\sqrt{2\pi}}
\]

The Fourier transform is a standard *Mathematica* command:
Simplify[FourierTransform[gauss[x, σ], x, ω], σ > 0]
FourierTransform[gauss[x, σ], x, ω]

A smaller kernel in the spatial domain gives a wider kernel in the Fourier domain, and vice versa.

Manipulate[
GraphicsRow[
{Plot[
\frac{e^{-\frac{x^2}{2\sigma^2}}}{\sigma \sqrt{2 \pi}}, \{x, -10, 10\}, PlotRange -> \{0, 0.4\},
PlotLabel -> "gauss[x, " <> ToString[σ] <> "]",
Plot[\(\mathcal{F}\text{gauss}[\omega, \sigma]\), \{\omega, -3, 3\}, PlotRange -> All,
PlotLabel -> "\(\mathcal{F}\text{gauss}[x, " <> ToString[σ] <> "]"\}], \{\{σ, 2\}, .5, 6, .2\}]
]

When applied to a signal, it operates as a lowpass filter.

scales = N[Table[\[FormalE]^t, \{t, 0, 8\}]]
spectra = (LogLinearPlot[\(\mathcal{F}\text{gauss}[\omega, \#1]\), \{\omega, 0.01`, 10\}] & @ scales);
Show[spectra, AspectRatio -> 0.4`, PlotRange -> All,
AxesLabel -> {\"\omega\", \"Amplitude\"}, ImageSize -> 300]

{1., 1.39561, 1.94773, 2.71828,
3.79367, 5.29449, 7.38906, 10.3123, 14.3919}
3.7 Central limit theorem

We see in the paragraph above the relation with the central limit theorem: any positive repetitive operator goes in the limit to a Gaussian function.

\[
\begin{align*}
\text{f[x_] := UnitStep[1/2 + x] + UnitStep[1/2 - x] - 1} & \quad \text{// Simplify;} \\
\text{g[x_] := UnitStep[1/2 + x] + UnitStep[1/2 - x] - 1} & \quad \text{// Simplify;} \\
\text{Plot[f[x], {x, -3, 3}, ImageSize -> 140]} & \\
\end{align*}
\]

Figure 3.3 The analytical blockfunction is a combination of two Heavyside unitstep functions.

We calculate analytically the convolution integral

\[
\text{h1 = Integrate[f[x] g[x - x1], \{x, -\infty, \infty\}]} \\
\quad \left\{ \begin{array}{l}
1 - x1 & 0 < x1 < 1 \\
1 + x1 & -1 < x1 \leq 0
\end{array} \right.
\]

\[
\text{Plot[h1, \{x1, -3, 3\}, PlotRange -> All, ImageSize -> 150]} \\
\]

Figure 3.4 One times a convolution of a blockfunction with the same blockfunction gives a triangle function.

The next convolution is this function convolved with the block function again:

\[
\text{h2 = Integrate[(h1 / . x1 -> x) g[x - x1], \{x, -\infty, \infty\}]} \\
\quad \left\{ \begin{array}{l}
\frac{1}{2} & x1 = \frac{1}{2} \\
\frac{1}{4} (3 - 4 x1^2) & -\frac{1}{2} < x1 < \frac{1}{2} \\
\frac{1}{8} (9 - 12 x1 + 4 x1^2) & \frac{1}{2} < x1 < \frac{3}{2} \\
\frac{1}{8} (9 + 12 x1 + 4 x1^2) & -\frac{3}{2} < x1 \leq -\frac{1}{2}
\end{array} \right.
\]

We see that we get a result that begins to look more towards a Gaussian:
3.7 Relation to binomial coefficients

\[ h(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \]

Figure 3.5 Two times a convolution of a block function with the same block function gives a function that rapidly begins to look like a Gaussian function. A Gaussian kernel with \( \sigma = 0.5 \) is drawn (dotted) for comparison.