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. This chapter discusses many of the attractive and special properties of the Gaussian kernel.

```
<< FrontEndVision`FEV`; Show[Import["Gauss10DM.gif"], ImageSize -> 280];
```

Figure 3.1 The Gaussian kernel is apparent on every German banknote of DM 10,- where it is depicted next to its famous inventor when he was 55 years old. The new Euro replaces these banknotes. See also: http://scienceworld.wolfram.com/biography/Gauss.html.

The Gaussian kernel is defined in 1-D, 2D and N-D respectively as

$$G_{1D}(x;\sigma)=\frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{x^2}{2\sigma^2}},\quad G_{2D}(x,y;\sigma)=\frac{1}{2\pi\sigma^2}\,e^{-\frac{x^2+y^2}{2\sigma^2}},\quad G_{ND}(\vec{x};\sigma)=\frac{1}{\left(\sqrt{2\pi}\,\sigma\right)^N}\,e^{-\frac{|\vec{x}|^2}{2\sigma^2}}$$

The σ determines the *width* of the Gaussian kernel. In statistics, when we consider the Gaussian probability density function it is called the *standard deviation*, and the square of it, σ^2, the *variance*. In the rest of this book, when we consider the Gaussian as an aperture function of some observation, we will refer to σ as the *inner scale* or shortly *scale*.

In the whole of this book the scale can only take positive values, $\sigma > 0$. In the process of observation σ can never become zero. For, this would imply making an observation through an infinitesimally small aperture, which is impossible. The factor of 2 in the exponent is a matter of convention, because we then have a 'cleaner' formula for the diffusion equation, as we will see later on. The semicolon between the spatial and scale parameters is conventionally put there to make the difference between these parameters explicit.

The scale-dimension is *not* just another spatial dimension, as we will thoroughly discuss in the remainder of this book.

The *half width at half maximum* ($\sigma = 2\sqrt{2 \ln 2}$) is often used to approximate σ, but it is somewhat larger:

```
Unprotect[gauss];
gauss[x_, σ_] := 1/(σ Sqrt[2 π]) Exp[-x^2/(2 σ^2)];
Solve[gauss[x, σ]/gauss[0, σ] == 1/2, x]
```

{{x → -σ √(2 Log[2])}, {x → σ √(2 Log[2])}}

```
% // N
```

{{x → -1.17741 σ}, {x → 1.17741 σ}}

3.2 Normalization

The term $\frac{1}{\sqrt{2\pi}\,\sigma}$ in front of the one-dimensional Gaussian kernel is the normalization constant. It comes from the fact that the integral over the exponential function is not unity: $\int_{-\infty}^{\infty} e^{-x^2/2\sigma^2}\,dx = \sqrt{2\pi}\,\sigma$. With the normalization constant this Gaussian kernel is a *normalized* kernel, i.e. its integral over its full domain is unity for every σ.

This means that increasing the σ of the kernel reduces the amplitude substantially. Let us look at the graphs of the normalized kernels for $\sigma = 0.3$, $\sigma = 1$ and $\sigma = 2$ plotted on the same axes:

```
Unprotect[gauss]; gauss[x_, σ_] := 1/(σ Sqrt[2 π]) Exp[-x^2/(2 σ^2)];

Block[{$DisplayFunction = Identity}, {p1, p2, p3} =
   Plot[gauss[x, σ = #], {x, -5, 5}, PlotRange -> {0, 1.4}] & /@
    {.3, 1, 2}];
Show[GraphicsArray[{p1, p2, p3}], ImageSize -> 400];
```

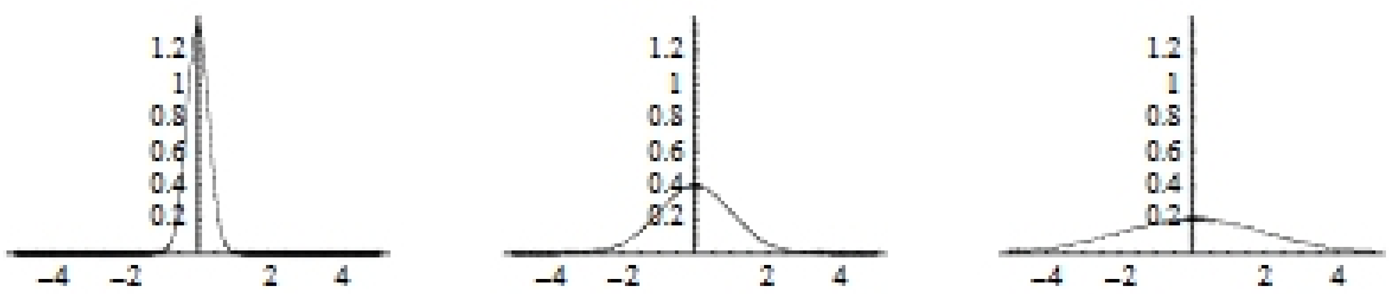

Figure 3.2 The Gaussian function at scales $\sigma = .3$, $\sigma = 1$ and $\sigma = 2$. The kernel is normalized, so the total area under the curve is always unity.

The normalization ensures that the average graylevel of the image remains the same when we blur the image with this kernel. This is known as *average grey level invariance*.

3.3 Cascade property, selfsimilarity

The shape of the kernel remains the same, irrespective of the σ. When we *convolve* two Gaussian kernels we get a new wider Gaussian with a variance σ^2 which is the sum of the variances of the constituting Gaussians: $g_{\text{new}}(\vec{x}; \sigma_1^2 + \sigma_2^2) = g_1(\vec{x}; \sigma_1^2) \otimes g_2(\vec{x}; \sigma_2^2)$.

```
σ =.; Simplify[∫_{-∞}^{∞} gauss[α, σ₁] gauss[α - x, σ₂] dα, {σ₁ > 0, σ₂ > 0}]
```

$$\frac{e^{-\frac{x^2}{2(\sigma_1^2+\sigma_2^2)}}}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}}$$

This phenomenon, i.e. that a new function emerges that is similar to the constituting functions, is called *self-similarity*.

The Gaussian is a *self-similar function*. Convolution with a Gaussian is a linear operation, so a convolution with a Gaussian kernel followed by a convolution with again a Gaussian kernel is equivalent to convolution with the broader kernel. Note that the *squares* of σ add, not the σ's themselves. Of course we can concatenate as many blurring steps as we want to create a larger blurring step. With analogy to a cascade of waterfalls spanning the same height as the total waterfall, this phenomenon is also known as the *cascade smoothing property*.
Famous examples of self-similar functions are *fractals*. This shows the famous Mandelbrot fractal:

```
cMandelbrot = Compile[{{c, _Complex}}, -Length[
    FixedPointList[#^2 + c &, c, 50, SameTest -> (Abs[#2] > 2.0 &)]]];
ListDensityPlot[-Table[cMandelbrot[a + b I], {b, -1.1, 1.1, 0.0114},
    {a, -2.0, 0.5, 0.0142}], Mesh -> False, AspectRatio -> Automatic,
  Frame -> False, ColorFunction -> Hue, ImageSize -> 170];
```

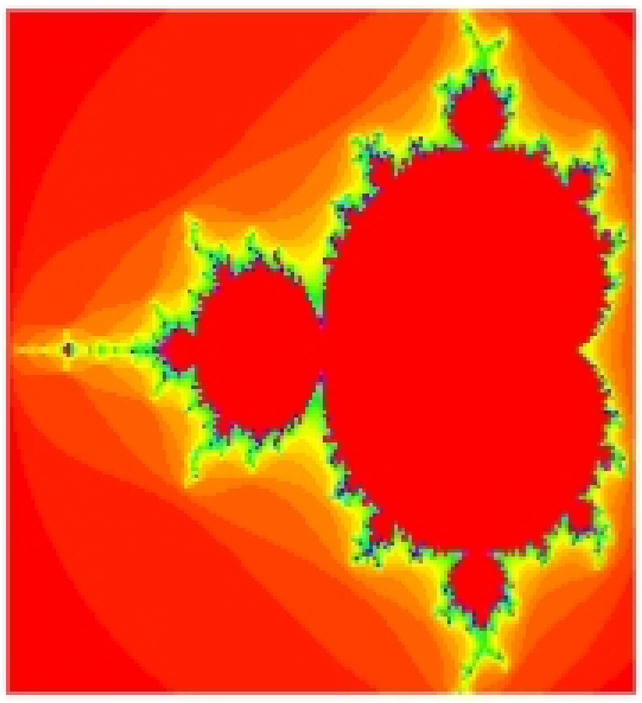

Figure 3.3 The Mandelbrot fractal is a famous example of a self-similar function. Source: www.mathforum.org. See also mathworld.wolfram.com/MandelbrotSet.html.