4. Gaussian derivatives

A difference which makes no difference is not a difference.
Mr. Spock (stardate 2822.3)

In[7]:=
```
url = "https://www.romeny.info/FEV-CD/images/";
```

4.1 Introduction

We will encounter the Gaussian derivative function at many places throughout this book. The Gaussian derivative function has many interesting properties. We will discuss them in one dimension first. We study its shape and algebraic structure, its Fourier transform, and its close relation to other functions like the Hermite functions, the Gabor functions and the generalized functions. In two and more dimensions additional properties are involved like orientation (directional derivatives) and anisotropy.

4.2 Shape and algebraic structure

When we take derivatives to *x* (*spatial derivatives*) of the Gaussian function repetitively, we see a pattern emerging of a polynomial of increasing order, multiplied with the original (normalized) Gaussian function again. Here we show a table of the derivatives from order 0 (i.e. no differentiation) to 3.

In[3]:=
```
gauss[x_, σ_] := 1/(σ Sqrt[2 π]) Exp[-x^2/(2 σ^2)];
Table[Factor[Evaluate[D[gauss[x, σ], {x, n}]]], {n, 0, 4}] // TableForm
```

Out[4]//TableForm=

$$\frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi}\,\sigma}$$

$$-\frac{e^{-\frac{x^2}{2\sigma^2}}\,x}{\sqrt{2\pi}\,\sigma^3}$$

$$\frac{e^{-\frac{x^2}{2\sigma^2}}\,(x-\sigma)\,(x+\sigma)}{\sqrt{2\pi}\,\sigma^5}$$

$$-\frac{e^{-\frac{x^2}{2\sigma^2}}\,x\,(x^2-3\sigma^2)}{\sqrt{2\pi}\,\sigma^7}$$

$$\frac{e^{-\frac{x^2}{2\sigma^2}}\,(x^4-6x^2\sigma^2+3\sigma^4)}{\sqrt{2\pi}\,\sigma^9}$$

The function **`Factor`** takes polynomial factors apart.

The function **`gauss[x,σ]`** is part of the standard set of functions (in **`FEV.m`**) with this book, and is protected. To modify it, it must be **`Unprotected`**.

The zeroth order derivative is indeed the Gaussian function itself. The even order (including the zeroth order) derivative functions are even functions (i.e. symmetric around zero) and the odd order derivatives are odd functions (antisymmetric around zero). This is how the graphs of Gaussian derivative functions look like, from order 0 up to order 7 (note the marked increase in amplitude for higher order of differentiation):

In[5]:=
```
GraphicsGrid[
 Partition[Table[Plot[Evaluate[D[gauss[x, 1], {x, n}]], {x, -5, 5},
    PlotLabel → StringJoin["Order=", ToString[n]]],
   {n, 0, 7}], 4], ImageSize → 500]
```

Out[5]=

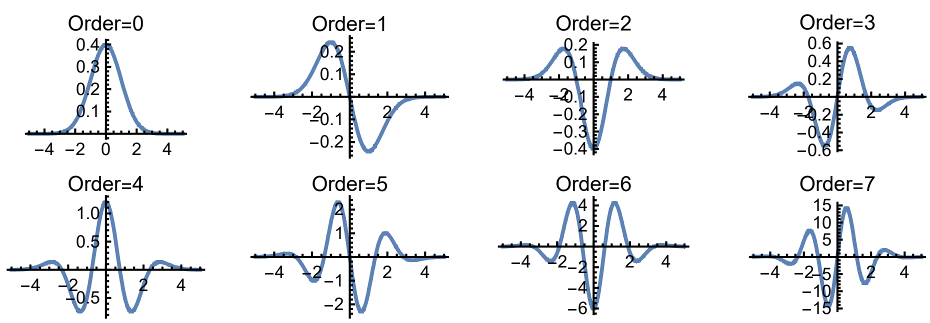

Figure 4.1 Plots of the 1D Gaussian derivative function for order 0 to 7.

The Gaussian function itself is a common element of all higher order derivatives. We extract the polynomials by dividing by the Gaussian function:

In[6]:=
```
Table[Evaluate[D[gauss[x, σ], {x, n}] / gauss[x, σ]], {n, 0, 4}] // Simplify
```

Out[6]=

$$\left\{1, -\frac{x}{\sigma^2}, \frac{x^2-\sigma^2}{\sigma^4}, -\frac{x^3-3x\sigma^2}{\sigma^6}, \frac{x^4-6x^2\sigma^2+3\sigma^4}{\sigma^8}\right\}$$

These polynomials have the same order as the derivative they are related to. Note that the highest order of x is the same as the order of differentiation, and that we have a plus sign for the highest order of x for even number of differentiation, and a minus signs for the odd orders.

These polynomials are the Hermite polynomials, called after Charles Hermite, a brilliant French mathematician (see figure 4.2).

In[9]:=
```
Import[url <> "CharlesHermite.jpg", ImageSize → 150]
```

Out[9]=

Figure 4.2 Charles Hermite (1822-1901).

They emerge from the following definition: $\frac{\partial^n e^{-x^2}}{\partial x^n} = (-1)^n\ H_n(x)\ e^{-x^2}$. The function $H_n(x)$ is the Hermite polynomial, where n is called the order of the polynomial. When we make the substitution $x \to x/(\sigma\ \sqrt{2})$, we get the following relation between the Gaussian function $G(x,\sigma)$ and its derivatives:
$\frac{\partial^n G(x,\sigma)}{\partial x^n} = (-1)^n \frac{1}{(\sigma\ \sqrt{2})^n}\ H_n\left(\frac{x}{\sigma\ \sqrt{2}}\right) G(x,\sigma)$.
In *Mathematica* the function H_n is given by the function **HermiteH[n,x]**. Here are the Hermite functions from zeroth to fifth order:

In[10]:=
```
Table[HermiteH[n, x], {n, 0, 7}] // TableForm
```

Out[10]//TableForm=

1
$2x$
$-2 + 4x^2$
$-12x + 8x^3$
$12 - 48x^2 + 16x^4$
$120x - 160x^3 + 32x^5$
$-120 + 720x^2 - 480x^4 + 64x^6$
$-1680x + 3360x^3 - 1344x^5 + 128x^7$

The inner scale σ is introduced in the equation by substituting $x \to \frac{x}{\sigma\ \sqrt{2}}$. As a consequence, with each differentiation we get a new factor $\frac{1}{\sigma\ \sqrt{2}}$. So now we are able to calculate the 1-D Gaussian derivative functions **gd[x,n,σ]** directly with the Hermite polynomials, again incorporating the normalization factor $\frac{1}{\sigma\ \sqrt{2\pi}}$:

In[11]:=
```
Clear[σ];
gd[x_, n_, σ_] := (-1/(σ Sqrt[2]))^n HermiteH[n, x/(σ Sqrt[2])] 1/(σ Sqrt[2 π]) Exp[-x^2/(2 σ^2)]
```

$$gd[x\_, n\_, \sigma\_] := \left(\frac{-1}{\sigma\ \sqrt{2}}\right)^n \text{HermiteH}\left[n, \frac{x}{\sigma\ \sqrt{2}}\right] \frac{1}{\sigma\ \sqrt{2\pi}} \text{Exp}\left[-\frac{x^2}{2\sigma^2}\right]$$

Check:

In[12]:=
```
Simplify[gd[x, 4, σ], σ > 0]
```

Out[12]=

$$\frac{e^{-\frac{x^2}{2\sigma^2}}\left(x^4-6x^2\sigma^2+3\sigma^4\right)}{\sqrt{2\pi}\,\sigma^9}$$

In[13]:=
$$\text{Simplify}\left[\text{D}\left[\frac{1}{\sigma\sqrt{2\pi}}\text{Exp}\left[-\frac{x^2}{2\sigma^2}\right],\{x,4\}\right],\sigma>0\right]$$

Out[13]=

$$\frac{e^{-\frac{x^2}{2\sigma^2}}\left(x^4-6x^2\sigma^2+3\sigma^4\right)}{\sqrt{2\pi}\,\sigma^9}$$

The amplitude of the Hermite polynomials explodes for large x, but the Gaussian envelop suppresses *any* polynomial function. No matter how high the polynomial order, the exponential function always wins. We can see this graphically when we look at e.g. the 7^{th} order Gaussian derivative without (i.e. the Hermite function, figure left) and with its Gaussian weight function (figure right). Note the vertical scales:

In[15]:=
$$f[x\_] := \left(\frac{1}{\sqrt{2}}\right)^7 \text{HermiteH}\left[7, \frac{x}{\sqrt{2}}\right];$$

$$\text{GraphicsRow}\left[\left\{\text{Plot}[f[x], \{x, -5, 5\}],\ p2 = \text{Plot}\left[f[x]\,\text{Exp}\left[-\frac{x^2}{2}\right], \{x, -5, 5\}\right]\right\}, \text{ImageSize} \to 400\right]$$

Out[15]=

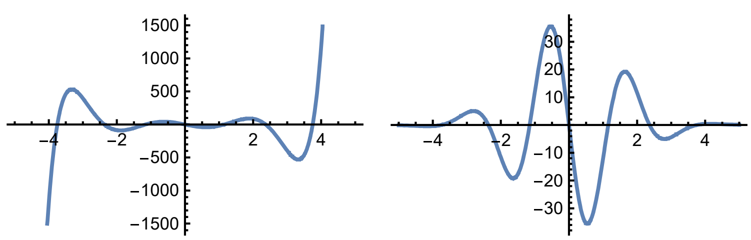

Figure 4.3 Left: The 7^{th} order Hermite polynomial. Right: idem, with a Gaussian envelop (weighting function). This is the 7^{th} order Gaussian derivative kernel.

Due to the limiting extent of the Gaussian window function, the amplitude of the Gaussian derivative function can be negligible at the location of the larger zeros. We plot an example, showing the 20^{th} order derivative and its Gaussian envelope function:

In[22]:=
```
n = 20;
σ = 1;
Plot[{gd[0, n, σ] × gauss[x, σ] / gauss[0, σ], gd[x, n, σ]}, {x, -5, 5}, ImageSize → 250]
```
Out[22]=

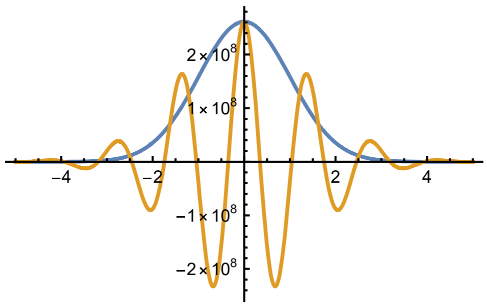

Figure 4.4 The 20th order Gaussian derivative's outer zero-crossings vahish in negligence. Note also that the amplitude of the Gaussian derivative function is not bounded by the Gaussian window. The Gabor kernels, as we will discuss later in section 4.7, are bounded by the Gaussian window.

How fast the Gaussian function goes zero can be seen from its values at $x = 3\,\sigma$, $x = 4\,\sigma$ and $x = 5\,\sigma$, relative to its peak value:

In[23]:=
```
Table[gauss[σ, 1] / gauss[0, 1], {σ, 3, 5}] // N
```
Out[23]=

$\{0.011109, 0.000335463, 3.72665 \times 10^{-6}\}$

and in the limit:

In[24]:=
```
Limit[gd[x, 7, 1], x → ∞]
```
Out[24]=

0

The Hermite polynomials belong to the family of orthogonal functions on the infinite interval (-∞,∞) with the weight function e^{-x^2}, $\int_{-\infty}^{\infty} e^{-x^2} H_n(x)\, H_m(x)\, dx = 2^{n+m}\, n!\, \sqrt{\pi}\, \delta_{nm}$, where δ_{nm} is the Kronecker delta, or delta tensor. $\delta_{nm} = 1$ for $n = m$, and $\delta_{nm} = 0$ for $n \neq m$.

In[25]:=
```
Table[∫_{-∞}^{∞} Exp[-x^2] HermiteH[k, x] HermiteH[m, x] dx,
  {k, 0, 3}, {m, 0, 3}] // MatrixForm
```
Out[25]//MatrixForm=

$$\begin{pmatrix} \sqrt{\pi} & 0 & 0 & 0 \\ 0 & 2\sqrt{\pi} & 0 & 0 \\ 0 & 0 & 8\sqrt{\pi} & 0 \\ 0 & 0 & 0 & 48\sqrt{\pi} \end{pmatrix}$$

The Gaussian derivative functions, with their weight function $e^{-\frac{x^2}{2}}$ are *not* orthogonal. We check this with some examples:

In[26]:=
```
{∫_{-∞}^{∞} gd[x, 2, 1] × gd[x, 3, 1] ⅆx, ∫_{-∞}^{∞} gd[x, 2, 1] × gd[x, 4, 1] ⅆx}
```

Out[26]=

$$\left\{0, -\frac{15}{16\sqrt{\pi}}\right\}$$

Other families of orthogonal polynomials are e.g. Legendre, Chebyshev, Laguerre, and Jacobi polynomials. Other orthogonal families of functions are e.g. Bessel functions and the spherical harmonic functions. The area under the Gaussian derivative functions is *not* unity, e.g. for the first derivative:

In[27]:=
```
SetOptions[Integrate, GenerateConditions → False];
∫_0^∞ gd[x, 1, σ] ⅆx
```

Out[28]=

$$-\frac{1}{\sqrt{2\pi}}$$

4.3 Gaussian derivatives in the Fourier domain

The Fourier transform of the derivative of a function is $(-i\,\omega)$ times the Fourier transform of the function. For each differentiation, a new factor $(-i\,\omega)$ is added. So the Fourier transforms of the Gaussian function and its first and second order derivatives are:

In[29]:=
```
σ =.;
Simplify[FourierTransform[
  {gauss[x, σ], ∂_x gauss[x, σ], ∂_{x,2} gauss[x, σ]}, x, ω], σ > 0]
```

Out[29]=

$$\left\{\frac{e^{-\frac{1}{2}\sigma^2\omega^2}}{\sqrt{2\pi}}, -\frac{i\, e^{-\frac{1}{2}\sigma^2\omega^2}\,\omega}{\sqrt{2\pi}}, -\frac{e^{-\frac{1}{2}\sigma^2\omega^2}\,\omega^2}{\sqrt{2\pi}}\right\}$$

In general: $\mathcal{F}\left\{\frac{\partial^n G(x,\sigma)}{\partial x^n}\right\} = (-\mathrm{i}\omega)^n\,\mathcal{F}\,\{G(x,\sigma)\}$.

Gaussian derivative kernels also act as bandpass filters. The maximum is at $\omega = \sqrt{n}$:

In[30]:=
```
n =.; σ = 1;
Solve[Evaluate[D[(-I ω)^n/Sqrt[2 π] Exp[-σ^2 ω^2/2], ω]] == 0, ω]
```

Out[31]=

$$\{\{\omega \to \sqrt{n}\}\}$$

The normalized powerspectra show that higher order of differentiation means a higher center frequency for the *bandpass* filter. The bandwidth remains virtually the same.

In[32]:=
```
σ = 1; Show[Table[Plot[Abs[((-I ω)^n/Sqrt[2 π] Exp[-σ^2 ω^2/2])/((-I n^(1/2))^n/Sqrt[2 π] Exp[-σ^2 n/2])], {ω, 0, 6}], {n, 1, 12}]]
```

Out[32]=

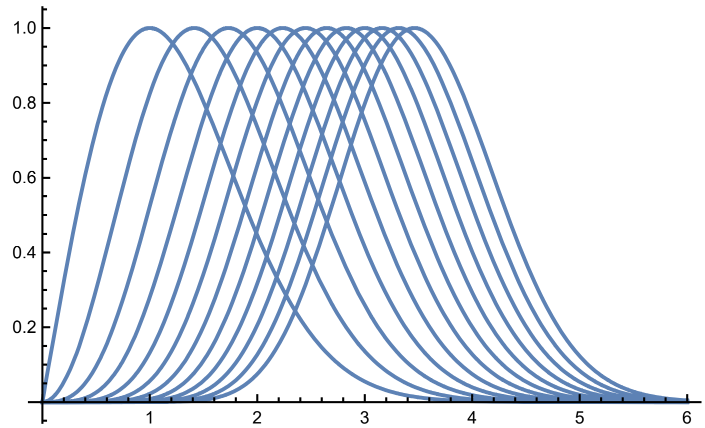

Figure 4.5 Normalized power spectra for Gaussian derivative filters for order 1 to 12, lowest order is left-most graph, $\sigma = 1$. Gaussian derivative kernels act like bandpass filters.

Task 4.0 Show with partial integration and the definitions from section 3.10 that the Fourier transform of the derivative of a function is $(-i\,\omega)$ times the Fourier transform of the function.

Task 4.0 Note that there are several definitions of the signs occurring in the Fourier transform (see the Wolfram Documentation in *Mathematica* under Fourier). Show that with the other definitions it is possible to arrive to the result that the Fourier transform of the derivative of a function is $(i\,\omega)$ times the Fourier transform of the function. In this book we stick to the default definition $(-i\,\omega)$.

4.4 Zero crossings of Gaussian derivative functions

In[36]:=

```
gd[x_, n_, σ_] := (-1/(σ Sqrt[2]))^n HermiteH[n, x/(σ Sqrt[2])] 1/(σ Sqrt[2 π]) Exp[-x^2/(2 σ^2)];
nmax = 20;
σ = 1;
Show[Graphics[Flatten[Table[{Blue, PointSize[0.01], Point [{n, x}]} /.
    Solve[HermiteH[n, x/Sqrt[2]] == 0, x], {n, 1, nmax}], 1]],
  AxesLabel → {"Order", "Zeros of\nHermiteH"}, Axes → True, ImageSize → 350]
```

Out[36]=

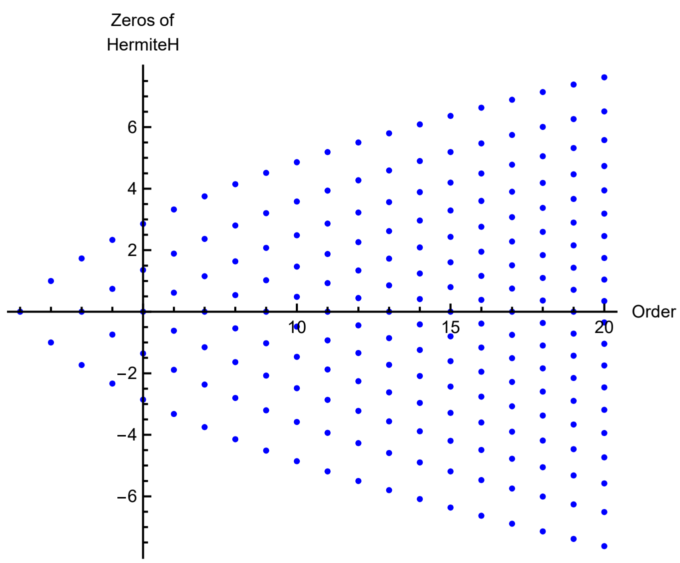

Figure 4.6 Zero crossings of Gaussian derivative functions to 20th order. Each dot is a zero-crossing.

How *wide* is a Gaussian derivative? This may seem a non-relevant question, because the Gaussian envelop often completely determines its behaviour. However, the number of zero-crossings is equal to the order of differentiation, because the Gaussian weighting function is a positive definite function.

It is of interest to study the behaviour of the zero-crossings. They move further apart with higher order. We can define the 'width' of a Gaussian derivative function as the distance between the outermost zero-crossings. The zero-crossings of the Hermite polynomials determine the zero-crossings of the Gaussian derivatives. In figure 4. 6all zeros of the first 20 Hermite functions as a function of the order are shown. Note that the zeros of the second derivative are just one standard deviation from the origin:

In[37]:=
```
σ =.; Simplify[Solve[D[gauss[x, σ], {x, 2}] == 0, x], σ > 0]
```
Out[37]=
```
{{x → -σ}, {x → σ}}
```

An exact analytic solution for the largest zero is not known. The formula of Zernicke (1931) specifies a range, and Szego (1939) gives a better estimate:

In[42]:=
```
p1 = Plot[2 Sqrt[n + 1 - 3.05 ∛(n + 1)], {n, 5, 50}];
(* Zernicke upper limit *)
p2 = Plot[2 Sqrt[n + 1 - 1.15 ∛(n + 1)], {n, 1, 50}];
(* Zernicke lower limit *)
p3 = Plot[2 √(n + .5) - 2.338098 / ⁶√(n + .5),
   {n, 1, 50}, PlotStyle → Dashing[{.01, .02}]];
Show[{p1, p2, p3}, AxesLabel →
  {"Order", "Width of Gaussian\nderivative (in σ)"}, ImageSize → 300]
```
Out[45]=

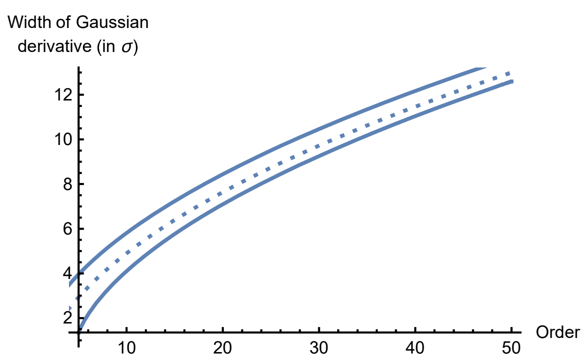

Figure 4.7 Estimates for the width of Gaussian derivative functions to 50th order. Width is defined as the distance between the outmost zero-crossings. Top and bottom graph: estimated range by Zernicke (1931), dashed graph: estimate by Szego (1939).

For very high orders of differentiation of course the numbers of zero-crossings increases, but also their mutual distance between the zeros becomes more equal. In the limiting case of infinite order the Gaussian derivative function becomes a sinusoidal function:

$$\lim_{n\to\infty} \frac{\partial^n G}{\partial^n x}(x, \sigma) = \mathrm{Sin}\left(x \sqrt{\frac{1}{\sigma}\left(\frac{n+1}{2}\right)}\right).$$

4.5 The correlation between Gaussian derivatives

Higher order Gaussian derivative kernels tend to become more and more similar. This makes them not very suitable as a basis. But before we investigate their role in a possible basis, let us investigate their similarity.

In fact we can express exactly how much they resemble each other as a function of the difference in differential order, by calculating the *correlation* between them. We derive the correlation below, and will appreciate the nice mathematical properties of the Gaussian function. Because the higher dimensional Gaussians are just the product of 1D Gaussian functions, it suffices to study the 1D case.

Compare e.g. the 20th and 24nd derivative function:

In[46]:=
```
g1 = Plot[gd[x, 20, 2], {x, -7, 7}, PlotLabel → "Order 20"];
g2 = Plot[gd[x, 24, 2], {x, -7, 7}, PlotLabel → "Order 24"];
GraphicsRow[{g1, g2}, ImageSize → 400]
```

Out[48]=

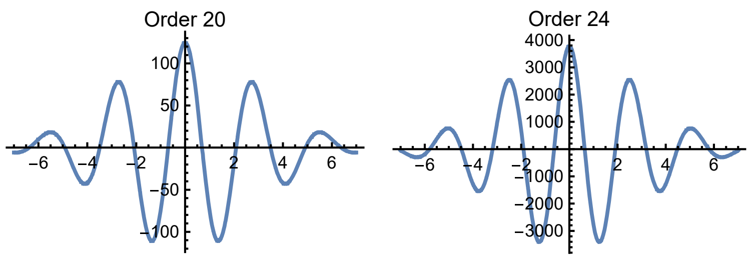

Figure 4.8 Gaussian derivative functions start to look more and more alike for higher order. Here the graphs are shown for the 20th and 24th order of differentiation.

The correlation coefficient between two functions is defined as the integral of the product of the functions over the full domain (in this case $-\infty$ to $+\infty$).

Because we want the coefficient to be unity for complete correlation (when the functions are identical by an amplitude scaling factor) we divide the coefficient by the so-called autocorrelation coefficients, i.e. the correlation of the functions with themselves.

We then get as definition for the correlation coefficient *r* between two Gaussian derivatives of order *n* and *m*:

$$r_{n,m} = \frac{\int_{-\infty}^{\infty} g^{(n)}(x)\, g^{(m)}(x)\, dx}{\sqrt{\int_{-\infty}^{\infty} \left[g^{(n)}(x)\right]^2 dx \int_{-\infty}^{\infty} \left[g^{(m)}(x)\right]^2 dx}}$$

with $g^{(n)}(x) = \frac{\partial^n g(x)}{\partial x^n}$. The Gaussian kernel $g(x)$ itself is an even function, and, as we have seen before, $g^{(n)}(x)$ is an even function for n is even, and an odd function for n is odd. The correlation between an even function and an odd function is zero. This is the case when n and m are both not even or both not odd, i.e. when $(n-m)$ is odd. We now can see already two important results:

$r_{n,m} = 0$ for $(n-m)$ odd;

$r_{n,m} = 1$ for $n = m$.

The remaining case is when $(n-m)$ is even. We take $n > m$. Let us first look to the nominator, $\int_{-\infty}^{\infty} g^{(n)}(x)\, g^{(m)}(x)\, dx$. The standard approach to tackle high exponents of functions in integrals, is the reduction of these exponents by partial integration:

$$\int_{-\infty}^{\infty} g^{(n)}(x)\, g^{(m)}(x)\, dx = g^{(n-1)}(x)\, g^{(m)}(x)\, |_{-\infty}^{\infty} - \int_{-\infty}^{\infty} g^{(n-1)}(x)\, g^{(m+1)}(x)\, dx = (-1)^k \int_{-\infty}^{\infty} g^{(n-k)}(x)\, g^{(m+k)}(x)\, dx$$

when we do the partial integration k times. The 'stick expression' $g^{(n-1)}(x)\, g^{(m)}(x)\, |_{-\infty}^{\infty}$ is zero because any Gaussian derivative function goes to zero for large x. We can choose k such that the exponents in the integral are equal (so we end up with the square of a Gaussian derivative function). So we make $(n-k) = (m+k)$, i.e. $k = \frac{(n-m)}{2}$. Because we study the case that $(n-m)$ is even, k is an integer number. We then get:

$$(-1)^k \int_{-\infty}^{\infty} g^{(n-k)}(x)\, g^{(m+k)}(x)\, dx = (-1)^{\frac{n-m}{2}} \int_{-\infty}^{\infty} g^{\left(\frac{n+m}{2}\right)}(x)\, g^{\left(\frac{n+m}{2}\right)}(x)\, dx$$

The *total energy* of a function in the spatial domain is the integral of the square of the function over its full extent. The famous theorem of Parceval states that the total energy of a function in the spatial domain is equal to the total energy of the function in the Fourier domain, i.e. expressed as the integral of the square of the Fourier transform over its full extent. Therefore

$$(-1)^{\frac{n-m}{2}} \int_{-\infty}^{\infty} g^{\left(\frac{n+m}{2}\right)}(x)\, g^{\left(\frac{n+m}{2}\right)}(x)\, dx \overset{\text{Parceval}}{=} (-1)^{\frac{n-m}{2}} \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| (i\omega)^{\frac{n+m}{2}}\, \hat{g}(\omega) \right|^2 d\omega =$$
$$(-1)^{\frac{n-m}{2}} \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^{n+m}\, \hat{g}^2(\omega)\, d\omega = (-1)^{\frac{n-m}{2}} \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^{n+m}\, e^{-\sigma^2 \omega^2}\, d\omega$$

We now substitute $\omega' = \sigma\, \omega$, and get finally: $(-1)^{\frac{n-m}{2}} \frac{1}{2\pi} \frac{1}{\sigma^{n+m+1}} \int_{-\infty}^{\infty} \omega'^{(n+m)}\, e^{-\omega'^2}\, d\omega'$.

This integral can be looked up in a table of integrals, but why not let *Mathematica* do the job (we first clear *n* and *m*):

In[49]:=
```
Clear[n, m]; ∫_{-∞}^{∞} x^(m+n) e^(-x^2) dx
```

Out[49]=

$$\frac{1}{2} \left(1 + (-1)^{m+n}\right) \text{Gamma}\left[\frac{1}{2}(1+m+n)\right] \text{Log}[e]^{\frac{1}{2}(-1-m-n)}$$

The function **Gamma** is the Euler gamma function. In our case **Re[m+n]>-1**, so we get for our correlation coefficient for (*n* – *m*) even:

$$r_{n,m} = \frac{(-1)^{\frac{n-m}{2}} \frac{1}{2\pi} \frac{1}{\sigma^{n+m+1}} \Gamma\left(\frac{m+n+1}{2}\right)}{\sqrt{\frac{1}{2\pi} \frac{1}{\sigma^{2n+1}} \Gamma\left(\frac{2n+1}{2}\right) \frac{1}{2\pi} \frac{1}{\sigma^{2m+1}} \Gamma\left(\frac{2m+1}{2}\right)}} = \frac{(-1)^{\frac{n-m}{2}} \Gamma\left(\frac{m+n+1}{2}\right)}{\sqrt{\Gamma\left(\frac{2n+1}{2}\right) \Gamma\left(\frac{2m+1}{2}\right)}}$$

Let's first have a look at this function for a range of values for n and m (0-15):

In[50]:=
```
r[n_, m_] := (-1)^((n-m)/2) Gamma[(m + n + 1)/2] / Sqrt[Gamma[(2 n + 1)/2] Gamma[(2 m + 1)/2]];
ListPlot3D[Table[Abs[r[n, m]], {n, 0, 15}, {m, 0, 15}],
 Axes → True, AxesLabel → {"n", "m", "Abs\nr[n,m]"},
 ViewPoint → {-2.348, -1.540, 1.281}, ImageSize → 220]
```

Out[51]=

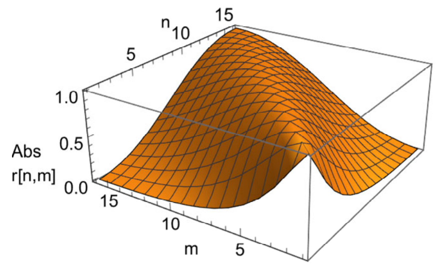

Figure 4.9 The magnitude of the correlation coefficient of Gaussian derivative functions for $0 < n < 15$ and $0 < m < 15$. The origin is in front.

Here is the function tabulated:.

In[52]:=
```
Table[NumberForm[r[n, m] // N, 3], {n, 0, 4}, {m, 0, 4}] // MatrixForm
```

Out[52]//MatrixForm=

1.	0. − 0.798 i	−0.577	0. + 0.412 i	0.293
0. + 0.798 i	1.	0. − 0.921 i	−0.775	0. + 0.623 i
−0.577	0. + 0.921 i	1.	0. − 0.952 i	−0.845
0. − 0.412 i	−0.775	0. + 0.952 i	1.	0. − 0.965 i
0.293	0. − 0.623 i	−0.845	0. + 0.965 i	1.

The correlation is unity when $n = m$, as expected, is negative when $n - m = 2$, and is positive when $n - m = 4$, and is complex otherwise. Indeed we see that when $n - m = 2$ the functions are even but of opposite sign:

In[53]:=
```
p1 = Plot[gd[x, 20, 2], {x, -5, 5}, PlotLabel → "Order 20"];
p2 = Plot[gd[x, 22, 2], {x, -5, 5}, PlotLabel → "Order 22"];
GraphicsRow[{p1, p2}, ImageSize → 450]
```

Out[55]=

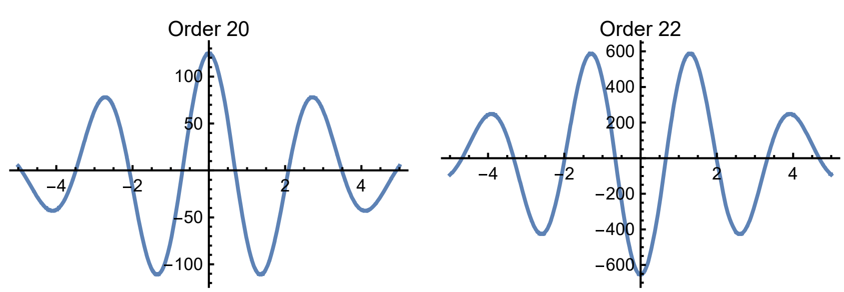

Figure 4.10 Gaussian derivative functions differing two orders are of opposite polarity.

and when $n - m = 1$ they have a phase-shift, leading to a complex correlation coefficient:

In[56]:=
```
p1 = Plot[gd[x, 20, 2], {x, -5, 5}, PlotLabel → "Order 20"];
p2 = Plot[gd[x, 21, 2], {x, -5, 5}, PlotLabel → "Order 21"];
GraphicsRow[{p1, p2}, ImageSize → 450]
```

Out[57]=

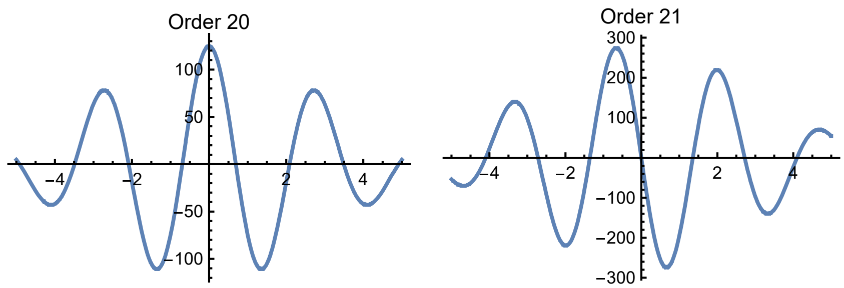

Figure 4.11 Gaussian derivative functions differing one order display a phase shift.

Of course, this is easy understood if we realize the factor $(-i\,\omega)$ in the Fourier domain, and that $i = e^{-i\frac{\pi}{2}}$. We plot the behaviour of the correlation coefficient of two close orders for large n. The asymptotic behaviour towards unity for increasing order is clear.

In[60]:=
```
Plot[-r[n, n + 2], {n, 1, 20}, AspectRatio → .4, PlotRange → {.8, 1.01},
 AxesLabel → {"Order", "Correlation\ncoefficient"},
 Epilog → {Dashed, Line[{{0, 1}, {20, 1}}]}]
```

Out[60]=

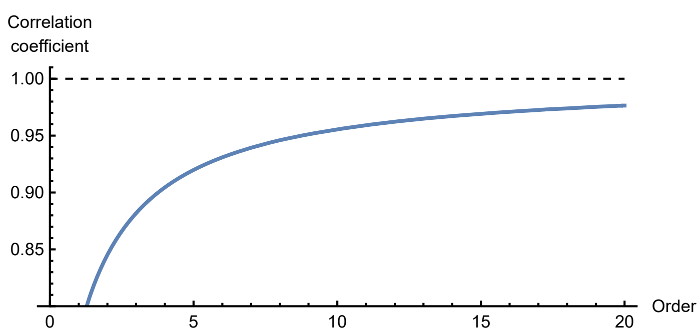

Figure 4.12 The correlation coefficient between a Gaussian derivative function and its even neighbour tends to unity for high differential order.

4.6 Discrete Gaussian kernels

In[63]:=
```
σ = 2;
Plot[{1/Sqrt[2 π σ^2] Exp[-x^2/(2 σ^2)], 1/Sqrt[2 π σ^2] BesselI[x, σ^2] / BesselI[0, σ^2]}, {x, 0,
  8}, PlotLegends → {"Gauss", "Bessel"}, PlotRange → All, ImageSize → 300]
```

Out[63]=

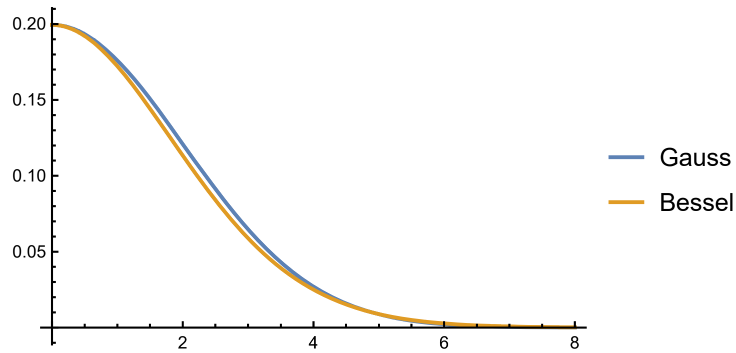

Figure 4.13 The graphs of the Gaussian kernel and the modified Bessel function of the first kind are very alike.

Lindeberg [Lindeberg1990] derived the optimal kernel for the case when the Gaussian kernel was discretized and came up with the "modified Bessel function of the first kind". In *Mathematica* this function is available as **`BesselI`**. This function is almost equal to the Gaussian kernel for $\sigma > 1$, as we see in figure 4.13. Note that the Bessel function has to be normalized by its value at $x = 0$. For larger σ the kernels become rapidly very similar.

4.7 Other families of kernels

The first principles that we applied to derive the Gaussian kernel in chapter 2 essentially stated "we know nothing" (at this stage of the observation). Of course, we can relax these principles, and introduce some knowledge. When we want to derive a set of apertures tuned to a *specific spatial frequency* $\vec{k}$ in the image, we add this physical quantity to the matrix of the dimensionality analysis:

In[64]:=
```
m = {{1, -1, -2, -2, -1}, {0, 0, 1, 1, 0}};
TableForm[m,
 TableHeadings → {{"meter", "candela"}, {"σ", "ω", "L0", "L", "k"}}]
```

Out[65]//TableForm=

	σ	ω	L0	L	k
meter	1	-1	-2	-2	-1
candela	0	0	1	1	0

The nullspace is now:

In[66]:=
```
NullSpace[m]
```

Out[66]=
```
{{1, 0, 0, 0, 1}, {0, 0, -1, 1, 0}, {1, 1, 0, 0, 0}}
```

Following the exactly similar line of reasoning, we end up from this new set of constraints with a new family of kernels, the *Gabor family of receptive fields*, with are given by a sinusoidal function (at the specified spatial frequency) under a Gaussian window.

In the Fourier domain: $Gabor(\omega, \sigma, k) = e^{-\omega^2 \sigma^2} e^{i k \omega}$, which translates into the spatial domain:

In[67]:=
$$\text{gabor}[\text{x\_}, \sigma\text{\_}] := \text{Sin}[\text{x}] \frac{1}{\sqrt{2 \pi \sigma^2}} \text{Exp}\left[-\frac{\text{x}^2}{2 \sigma^2}\right];$$

The Gabor function model of cortical receptive fields was first proposed by Marcelja in 1980 [Marcelja1980]. However the functions themselves are often credited to Gabor [Gabor1946] who supported their use in communications.

Gabor functions are defined as the sinus function under a Gaussian window with scale σ. The phase ϕ of the sinusoidal function determines its detailed behaviour, e.g. for $\phi = \pi/2$ we get an even function.Gabor functions can look very much like Gaussian derivatives, but there are essential differences:

- Gabor functions have an infinite number of zero-crossings on their domain.

- The amplitudes of the sinusoidal function never exceeds the Gaussian envelope.

In[93]:=
```
gabor[x_, ϕ_, σ_] := Sin[x + ϕ] gauss[x, σ];
GraphicsRow[{Plot[{gabor[x, 0, 10], gauss[x, 10]}, {x, -30, 30},
   PlotRange → {-.04, .04}], Plot[{gabor[x, π / 2, 10], gauss[x, 10]},
   {x, -30, 30}, PlotRange → {-.04, .04}]}, ImageSize → 500]
```

Out[94]=

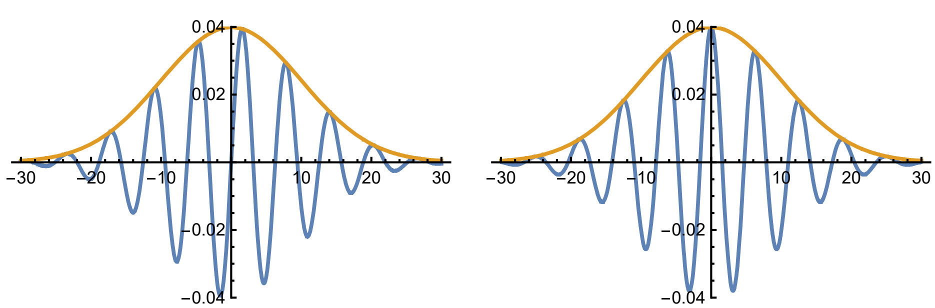

Figure 4.14 Gabor functions are sinusoidal functions with a Gaussian envelope. Left: Sin[x] G[x,10]; right: Sin[x+π/2] G[x,10].

Gabor functions can be made to look very similar by an appropriate choice of parameters:

In[107]:=

```
σ = 1; gd[x_, σ_] = D[1/Sqrt[2 π σ^2] Exp[-x^2/(2 σ^2)], x];
Plot[{- 1.2 gabor[x, 1.2], gd[x, 1]}, {x, -4, 4},
PlotLegends → LineLegend[{"Gabor", "Gauss"}, LegendLabel → "Function:",
   LegendFunction → (Framed[#, RoundingRadius → 5] &),
   LegendMargins → 5], ImageSize → 300]
```

Out[108]=

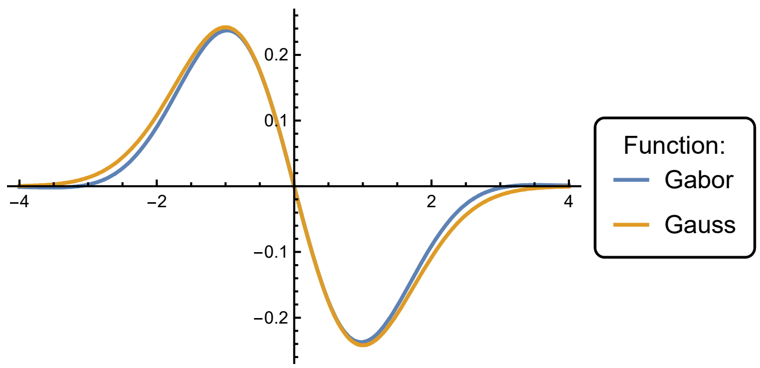

Figure 4.15 Gabor functions can be made very similar to Gaussian derivative kernels. In a practical application then there is no difference in result. Dotted graph: Gaussian first derivative kernel. Continuous graph: Minus the Gabor kernel with the same σ as the Gaussian kernel. Note the necessity of sign change due to the polarity of the sinusoidal function.

If we relax one or more of the first principles (leave one or more out, or add other axioms), we get other families of kernels. E.g. when we add the constraint that the kernel should be tuned to a specific spatial frequency, we get the family of *Gabor kernels* [Florack1992a, Florack1997a]. It was recently shown by Duits et al. [Duits2002a], extending the work of Pauwels [Pauwels1995], that giving up the constraint of separability gives a new family of interesting *Poisson* scale-space kernels, defined by the solution of the Dirichlet problem $\frac{\partial L}{\partial s} = -(-\Delta)^{\alpha} L$. For $\alpha = 1$ we find the Gaussian scale-space, for $\alpha = \frac{1}{2}$ we get the Poisson scale-space. In this book we limit ourselves to the Gaussian kernel.

We conclude this section by the realization that the front-end visual system at the retinal level must be uncommitted, no feedback from higher levels is at stake, so the Gaussian kernel seems a good candidate to start observing with at this level. At higher levels this constraint is released.

The extensive feedback loops from the primary visual cortex to the LGN may give rise to 'geometry-driven diffusion' [TerHaarRomeny1994f], nonlinear scale-space theory, where the early differential geometric measurements through e.g. the simple cells may modify the kernels LGN levels. Nonlinear scale-space theory will be treated in chapter 21.

Task 4.2 When we have noise in the signal to be differentiated, we have two counterbalancing effect when we change differential order and scale: for higher order the noise is amplified (the factor $(-i\,\omega)^n$ in the Fourier transform representation) and the noise is averaged out for larger scales. Give an explicit formula in our *Mathematica* framework for the propagation of noise when filtered with Gaussian derivatives. Start with the easiest case, i.e. pixel-uncorrelated (white) noise, and continue with correlated noise. See for a treatment of this subject the work by Blom et al. [Blom1993a].

Task 4.3 Give an explicit formula in our *Mathematica* framework for the propagation of noise when filtered with a *compound function* of Gaussian derivatives, e.g. by the Laplacian $\frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2}$. See for a treatment of this subject the work by Blom et al. [Blom1993a].

4.8 Higher dimensions and separability

Gaussian derivative kernels of higher dimensions are simply made by multiplication. Here again we see the separability of the Gaussian, i.e. this *is* the separability. The function **gd2D[x,y,n,m,σx,σy]** is an example of a Gaussian partial derivative function in 2D, first order derivative to x, second order derivative to y, at scale 2 (equal for x and y):

In[109]:=

```
gd2D[x_, y_, n_, m_, σx_, σy_] := gd[x, n, σx] × gd[y, m, σy];
Plot3D[gd2D[x, y, 1, 2, 2, 2], {x, -7, 7},
 {y, -7, 7}, AxesLabel → {x, y, ""}, PlotPoints → 40,
 PlotRange → All, Boxed → False, Axes → True, ImageSize → 300]
```

Out[110]=

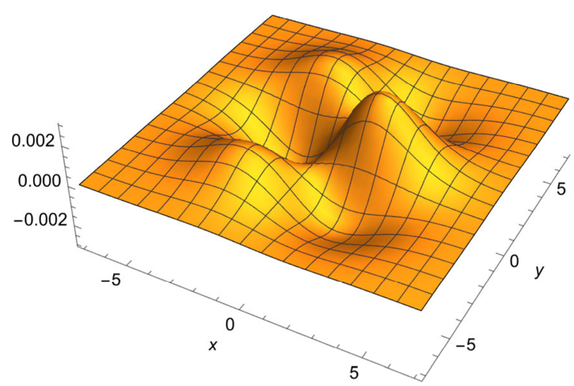

Figure 4.16 Plot of $\frac{\partial^3 G(x,y)}{\partial x\,\partial y^2}$. The two-dimensional Gaussian derivative function can be constructed as the product of two one-dimensional Gaussian derivative functions, and so for higher dimensions, due to the separability of the Gaussian kernel for higher dimensions.

The ratio $\frac{\sigma_x}{\sigma_y}$ is called the anisotropy ratio. When it is unity, we have an *isotropic* kernel, which diffuses in the *x* and *y* direction by the same amount. The Greek word 'isos' ($\iota\sigma o\varsigma$) means 'equal', the Greek word 'tropos' ($\tau\rho o\pi o\varsigma$) means 'direction' (the Greek word 'topos' ($\tau o\pi o\varsigma$) means 'location, place').

In 3D the iso-intensity surfaces of the Gaussian kernels are shown (and can be interactively manipulated) with the command MVContourPlot3D from the OpenGL

viewer 'MathGL3D' by J.P.Kuska
(phong.informatik.uni-leipzig.de/~kuska/mathgl3dv3):

```
set = D[E^(-(x^2+y^2+z^2)/(2 σ^2)), {x, #}] & /@ {1, 2, 3};
GraphicsRow[
 ContourPlot3D[#, {x, -3.5, 3.5}, {y, -3.5, 3.5}, {z, -3, 0}, Contours →
     Range[-.6, .6, .1], BoxRatios → {2, 2, 1}] & /@ set, ImageSize → 500]
```

Out[123]=

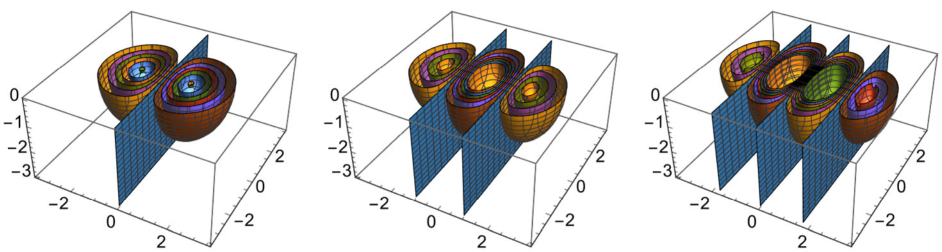

Figure 4.17 Iso-intensity surfaces for Gaussian derivative kernels in 3D. Left: $\frac{\partial G}{\partial x}$; middle: $\frac{\partial^2 G}{\partial x^2}$; right: $\frac{\partial^3 G}{\partial x^3}$.

The sum of 2 of the three 2nd order derivatives is called the 'hotdog' detector:

In[124]:=

```
σ = 1;
f = Evaluate[∂x,x E^(-(x^2+y^2+z^2)/(2 σ^2)) + ∂z,z E^(-(x^2+y^2+z^2)/(2 σ^2))];
ContourPlot3D[f, {x, -3.5, 3.5}, {y, -3.5, 3.5}, {z, -3, 0},
 Contours → Range[-.6, .6, .1], BoxRatios → {2, 2, 1}, ImageSize → 250]
```

Out[125]=

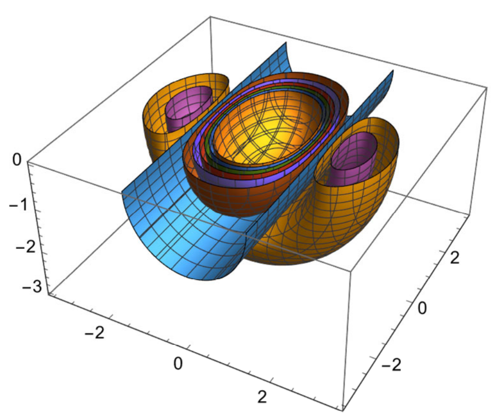

Figure 4.18 Iso-intensity surface for the Gaussian derivative in 3D $\frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial z^2}$.

4.9 Summary of this chapter

The Gaussian derivatives are characterized by the product of a polynomial function, the Hermite polynomial, and a Gaussian kernel. The order of the Hermite polynomial is the same as the differential order of the Gaussian derivative. Many interesting recursive relations exist for Hermite polynomials, making them very suitable for analytical treatment. The shape of the Gaussian derivative kernel can be very similar to specific Gabor kernels. One essential difference is the number of zerocrossing: this is always infinite for Gabor kernels, the number of zerocrossings of Gaussian derivatives is equal to the differential order. The envelope of the Gaussian derivative amplitude is not the Gaussian function, as is the case for Gabor kernels.

The even orders are symmetric kernels, the odd orders are asymmetric kernels. The normalized zeroth order kernel has unit area by definition, the Gaussian derivative kernels of the normalized kernel have no unit area.

Gaussian derivatives are not orthogonal kernels. They become more and more correlated for higher order, if odd or even specimens are compared. The limiting case for infinite order leads to a sinusoidal (for the odd orders) or cosinusoidal (for the even orders) function with a Gaussian envelope, i.e. a Gabor function.

In the vision chapters we will encounter the Gaussian derivative functions as suitable and likely candidates for the receptive fields in the primary visual cortex.