

## The *error function* as an integration trick

### Abstract

The *error function*  $\text{erf}(x)$  is a statistical function but can be used as an integration trick for complex exponentials, which occur frequently in physics, for example in the differential equations of heat propagation. It is also used in quantum mechanics to describe particles represented by a wave function, or in astrophysics for the spectroscopic analysis of spectral lines, as in the examples discussed in detail.

### Definition and statistical meaning

The *error function* is defined by the integral:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt . \quad (1)$$

Its statistical meaning is to represent the probability that a random variable with a normal distribution (i.e. Gaussian, with mean 0 and variance 1/2) falls in the interval  $[-x, x]$ . It has notable values:

$$\text{erf}(0) = 0 , \text{erf}(\infty) = 1 , \text{erf}(-\infty) = -1 . \quad (2)$$

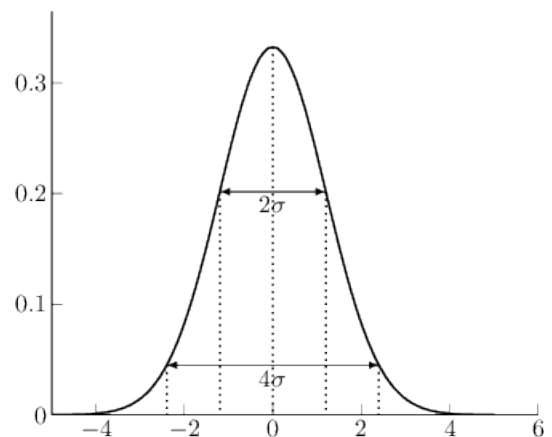
Of course  $2/\sqrt{\pi}$  represents the normalization factor, which ensures that the total probability is 1, while the function determines the area included between the Gaussian distribution curve included in  $[-x, x]$ .

The *probability density function* (bell-shaped curve, see Figure 1), which gives the value at the point  $x$ , is defined by the *classical Gaussian equation*:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} , \quad (3)$$

where  $\mu$  It is the mean that determines the center of the bell and  $\sigma$  is the standard deviation that determines the width of the bell.

Figure 1 - Gaussian curve



The *normal distribution function*  $F(x)$  represents the total area to the left of a point  $x$ , or the probability that a value is less than or equal to  $x$ , and is defined on the basis of (3) as:

$$F(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt . \quad (4)$$

The (3) formula can be simplified into the so-called standard formulation (usually indicated with the letter  $\Phi$ ), considering  $\mu = 0$  e  $\sigma = 1$ , getting:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} . \quad (5)$$

So we can link the error function to the probability of a standard normal distribution, with the formula, derived from (1) and (5) :

$$\Phi(x) = \frac{1}{2} \left[ 1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) \right] . \quad (6)$$

In other words, from (6) one can directly derive the probability that an event falls within  $x$  standard deviations from the mean, i.e.  $|Z| \leq x$ , using the function:

$$P(|Z| \leq x) = \Phi(x) - \Phi(-x) = 2\Phi(x) - 1 = \operatorname{erf}(x/\sqrt{2}) , \quad (7)$$

obtaining for example the classic values:

$$68.27 \% \text{ if } \sigma = 1 , 97.73 \% \text{ if } \sigma = 2 , 99.87 \% \text{ if } \sigma = 3 .$$

Use as integration trick

A classic example of integration using (1) is given by the integration of the Gaussian function, which can be solved as:

$$\int e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \operatorname{erf}(x) + \text{const} , \quad (8)$$

by transforming an unsolvable integral into a known function, which even calculators recognize.

Let's get to the trick we're going to explore in depth, the case of an integral of the type:

$$\int e^{-(x^2+2x)} dx , \quad (9)$$

which happens quite often in the study of physics.

Rewriting the exponent of (9) as  $-(x+1)^2 + 1$ , applying (8) the integral becomes:

$$e^1 \int e^{-(x+1)^2} dx = e \frac{\sqrt{\pi}}{2} \operatorname{erf}(x+1) + \text{const} , \quad (10)$$

possibly by substituting variables in more complex cases.

*Without the error function*, we couldn't write the result of this integral exactly.

And when the result happens to be part of a more complex (in the sense of complicated) function to analyze, we know that it is a finite number, if  $x \in \mathbb{R}$ , included in the interval  $[-1, 1]$ , except for coefficients that can scale the value.

Example of use in quantum mechanics

A free particle, such as an electron, is described by a wave function that can be imagined as Gaussian in shape. To describe its position, we must solve an integral of the form:

$$I = \int_{-\infty}^{+\infty} e^{-(ax^2 + bx)} dx, \quad (11)$$

where  $x$  is the position,  $a$  is related to the width of the wave packet by the relation  $\Delta x \approx 1/\sqrt{a}$  and  $b$  is related to the speed (at the moment) of the particle.

Applying the trick seen in (10), the exponent of (11) can be rewritten as:

$$(ax^2 + bx) = a \left( x^2 + \frac{b}{a} x \right), \quad (12)$$

then from (12) obtain a quadratic form:

$$a \left( x^2 + \frac{b}{a} x + \frac{b^2}{4a^2} - \frac{b^2}{4a^2} \right) = a \left( x + \frac{b}{2a} \right)^2 - \frac{b^2}{4a^2}. \quad (13)$$

Substituting (13) into equation (11), taking the constants out of the integration, we obtain:

$$I = e^{\frac{b^2}{4a^2}} \int_{-\infty}^{+\infty} e^{-a \left( x + \frac{b}{2a} \right)^2} dx. \quad (14)$$

By applying the *error function* we can directly integrate (14) which is a translated Gaussian, remembering the values (2), obtaining:

$$I = e^{\frac{b^2}{4a^2}} \sqrt{\frac{\pi}{a}}. \quad (15)$$

### Quantum note

The (15) formula suggests that in the case of a very narrow wave packet, i.e., a precisely known initial position, it will expand significantly over time, as predicted by *Heisenberg's uncertainty principle*.

Mathematically, to convert from the position to the velocity of the particle (the momentum), we use the Fourier transform and the probability density function related to the position:

$$\Psi(x) \propto e^{-ax^2}, \quad (16)$$

for the momentum  $p$ , is identified by the integral:

$$\Psi(p) = \int_{-\infty}^{+\infty} e^{-ax^2} e^{-ipx} dx. \quad (17)$$

By solving (17) with the trick of completing the square, we obtain a Gaussian in the variable  $p$ :

$$\Psi(p) \propto e^{-\frac{p^2}{4a}}. \quad (18)$$

Since the exponent is negative and the value of  $a$  it is assumed to be large (well defined position), then the probability density function (18) widens tending to zero very slowly (large dispersion therefore large imprecision of the velocity), consistently with the *uncertainty principle*.

Example of use for spectroscopic analysis of spectral lines

To perform the spectroscopic analysis of the spectral lines, which always have a certain width, a convolution of the *natural broadening* components (Lorentzian profile) and the *Doppler broadening* components (Gaussian profile) is considered, which leads to a function, called *Voigt profile* (see Figure 2), which is described by a rather complicated integral:

$$V(x; \sigma, \gamma) = \int_{-\infty}^{+\infty} G(t; \sigma) L(x - t; \gamma) dt$$

where  $x$  represents the distance from the center point,  $G(x; \sigma) \equiv \frac{e^{-x^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$  is the Gaussian centered profile and  $L(x; \gamma) \equiv \frac{\gamma}{\pi(\gamma^2 + x^2)}$  it is the Lorentzian centered profile.

Solving it is off-topic for this study, but with the trick of completing the square, we obtain that the spectral line is a Gaussian type curve centered on the shifted frequency, whose width depends on the temperature; for this reason, it is called *thermal broadening*.

Figure 2 - Spectral line broadening profiles

