

Fourier Transforms

For additional information, see the classic book *The Fourier Transform and its Applications* by Ronald N. Bracewell (which is on the shelves of most radio astronomers) and the [Wikipedia](#) and [Mathworld](#) entries for the Fourier transform.

The Fourier transform is important in mathematics, engineering, and the physical sciences. Its discrete counterpart, the Discrete Fourier Transform (DFT), which is normally computed using the so-called Fast Fourier Transform (FFT), has revolutionized modern society, as it is ubiquitous in digital electronics and signal processing. Radio astronomers are particularly avid users of Fourier transforms because Fourier transforms are key components in data processing (e.g., periodicity searches) and instruments (e.g., antennas, receivers, spectrometers), and they are the cornerstones of interferometry and aperture synthesis.

The Fourier transform is a reversible, linear transform with many important properties. For any function $f(x)$ (which in astronomy is usually real-valued, but $f(x)$ may be complex), the Fourier transform can be denoted $F(s)$, where the product of x and s is dimensionless. Often x is a measure of time t (i.e., the *time-domain* signal) and so s corresponds to inverse time, or frequency ν (i.e., the *frequency-domain* signal).

The **Fourier transform** is defined by

$$F(s) \equiv \int_{-\infty}^{\infty} f(x) e^{-2\pi i s x} dx, \quad (\text{SF1})$$

which is usually known as the **forward transform**, and

$$f(x) \equiv \int_{-\infty}^{\infty} F(s) e^{2\pi i s x} ds, \quad (\text{SF2})$$

which is the **inverse transform**. In both cases, $i \equiv \sqrt{-1}$. Alternative definitions of the Fourier transform are based on angular frequency ($\omega = 2\pi\nu$), have different normalizations, or the opposite sign convention in the complex exponential. Since Fourier transformation is reversible, the symmetric symbol \Leftrightarrow is often used to mean "is the Fourier transform of"; e.g., $F(s) \Leftrightarrow f(x)$.

The complex exponential is the heart of the transform. A complex exponential is simply a complex number where both the real and imaginary parts are sinusoids. The exact relation is called **Euler's formula**

$$e^{i\phi} = \cos \phi + i \sin \phi, \quad (\text{SF3})$$

which leads to the famous (and beautiful) identity $e^{i\pi} + 1 = 0$ that relates five of the most important numbers in mathematics. Complex exponentials are much easier to manipulate than trigonometric functions, and they provide a compact notation for dealing with sinusoids of arbitrary phase, which form the basis of the Fourier transform.

Complex exponentials (or sines and cosines) are periodic functions, and the set of complex exponentials is complete and orthogonal. Thus the Fourier transform can represent any piecewise continuous function and minimizes the least-square error between the function and its representation. There exist other complete and orthogonal sets of periodic functions; for example, [Walsh functions](#) (square waves) are useful for digital electronics. Why do we always encounter complex exponentials when solving physical problems? Why are monochromatic waves sinusoidal, and not periodic trains of square waves or triangular waves? The reason is that the derivatives of complex exponentials are just rescaled complex exponentials. In other words, the complex exponentials are the eigenfunctions of the differential operator. Most physical systems obey linear differential equations. Thus an analog electronic filter will convert a sine wave into another sine wave having the same frequency (but not necessarily the same amplitude and phase), while a filtered square wave will not be a square wave. This property of complex exponentials makes the Fourier transform uniquely useful in fields ranging from radio propagation to quantum mechanics.

The Discrete Fourier Transform

The continuous Fourier transform converts a time-domain signal of infinite duration into a continuous spectrum composed of an infinite number of sinusoids. In astronomical observations we deal with signals that are discretely sampled, usually at constant intervals, and of finite duration or periodic. For such data, only a finite number of sinusoids is needed and the **Discrete Fourier Transform** (DFT) is appropriate. For almost every Fourier transform theorem or property, there is a related theorem or property for the DFT. The DFT of N uniformly sampled data points x_j (where $j = 0, \dots, N-1$) and its inverse are defined by

$$X_k = \sum_{j=0}^{N-1} x_j e^{-2\pi i j k / N} \quad (\text{SF4})$$

and

$$x_j = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{2\pi i j k / N} . \quad (\text{SF5})$$

Once again, sign and normalization conventions may vary, but our definition is the most common. The continuous variable s has been replaced by the discrete variable (usually an integer) k .

The result of the DFT of an N -point input time series is an N -point frequency spectrum, with Fourier frequencies k ranging from $-(N/2 - 1)$, through the 0-frequency or so-called "DC" component, and up to the highest Fourier frequency $N/2$. Each bin number represents the integer number of sinusoidal periods present in the time series. The amplitudes and phases represent the amplitudes A_k and phases ϕ_k of those sinusoids. In summary, each bin can be described by $X_k = A_k e^{i\phi_k}$.

For real-valued input data, however, the resulting DFT is **hermitian**—the real-part of the spectrum is an even function and the imaginary part is odd, such that $X_{-k} = \overline{X_k}$, where the bar represents complex conjugation. This means that all of the "negative" Fourier frequencies provide no new information. Both the $k=0$ and $k=N/2$ bins are real-valued, and there is a total of $N/2+1$ Fourier bins,

so the total number of independent pieces of information (i.e. real and complex parts) is N , just as for the input time series. No information is created or destroyed by the DFT.

Other symmetries exist between time- and frequency-domain signals as well:

<u>Time Domain</u>	<u>Frequency Domain</u>
real	hermitian (real=even, imag=odd)
imaginary	anti-hermitian (real=odd, imag=even)
even	even
odd	odd
real and even	real and even (i.e. cosine transform)
real and odd	imaginary and odd (i.e. sine transform)
imaginary and even	imaginary and even
imaginary and odd	real and odd

Usually the DFT is computed by a very clever (and truly revolutionary) algorithm known as the **Fast Fourier Transform** or FFT. The FFT was discovered by Gauss in 1805 and re-discovered many times since, but most people attribute its modern incarnation to James W. Cooley and John W. Tukey ("An algorithm for the machine calculation of complex Fourier series," *Math. Comput.* **19**, 297–301) in 1965. The key advantage of the FFT over the DFT is that the operational complexity decreases from $O(N^2)$ for a DFT to $O(N \log_2(N))$ for the FFT. Modern implementations of the FFT (such as [FFTW](#)) allow $O(N \log_2(N))$ complexity for *any* value of N , not just those that are powers of two or the products of only small primes.

Example: Estimate the speed increases obtained by computing the FFTs instead of DFTs for transforms of length 10^3 , 10^6 , and 10^9 points.

$$\text{speed improvement } (N = 10^3) \propto \frac{N^2}{N \log_2(N)} = \frac{N}{\log_2(N)} \sim \frac{10^3}{10} \sim 100$$

$$\text{speed improvement } (N = 10^6) \propto \frac{N}{\log_2(N)} \sim \frac{10^6}{20} \sim 5 \times 10^4$$

$$\text{speed improvement } (N = 10^9) \propto \frac{N}{\log_2(N)} \sim \frac{10^9}{30} \sim 3 \times 10^7$$

Sampling Theorem

For a DFT to represent a function accurately, the original function must be sampled at a sufficiently high rate. The appropriate rate for a uniformly sampled time series is determined by the **Nyquist-Shannon Theorem** or **Sampling Theorem**. This theorem states that any continuous baseband signal (signal extending down to zero frequency) may be identically reconstructed if the signal is bandwidth limited and the sampling frequency is *at least* twice the bandwidth of the signal (i.e. the highest frequency of a baseband signal). That critical sampling rate, $1/\Delta t$, where Δt is the time between successive samples, is known as the **Nyquist rate**, and it is a property of the time-domain signal based on its frequency content. Somewhat confusingly, if a time-domain signal is sampled uniformly, then the frequency corresponding to one-half that rate is called the **Nyquist frequency**,

$$\nu_{N/2} = 1/(2 \Delta t) . \quad (\text{SF6})$$

The Nyquist frequency describes the high frequency cut-off of the system doing the sampling, and is therefore a property of that system. Any frequencies present in the original signal which are at higher frequencies than the Nyquist frequency will be **aliased** to other lower frequencies in the sampled band as described below. If that signal was band-limited and then sampled at the Nyquist rate, in accordance to the Sampling Theorem, no aliasing will occur.

In a DFT, where there are N samples spanning a total time $T = N \Delta t$, the frequency resolution is $1/T$. Each Fourier bin number k represents exactly k sinusoidal oscillations in the original data x_j , and therefore a frequency $\nu = k/T$ in Hz. The Nyquist frequency corresponds to bin $k = \nu_{N/2} T = T/(2 \Delta t) = NT/(2T) = N/2$. If the signal is not bandwidth limited and higher-frequency components [with $k > N/2$ or $\nu > N/(2T)$ Hz] exist, those frequencies will show up in the DFT aliased back into lower frequencies $f_a = N/T - \nu$, assuming that $N/(2T) < \nu < N/T$. Such aliasing can be avoided by filtering the input data to ensure that it is properly **band-limited**.

Examples:

The (young and undamaged) human ear can hear sounds with frequency components up to around 20 kHz. Therefore, near-perfect audio recording systems must sample audio signals at at least 40 kHz to be Nyquist sampled. Because of the practical nature of real-life audio filters (which do not have infinitely sharp cutoffs), audio CDs are sampled at 44100 Hz. That gives a 2 kHz buffer which allows imperfect lowpass audio filters to filter out higher frequencies which would otherwise be aliased into the audible band.

A visual example of an aliased signal often occurs in western movies where the 24 frame-per-second rate of the movie camera performs "stroboscopic" sampling of a rapidly spinning wagon wheel. When the rotation rate of the wheel is below the Nyquist rate (12 Hz), it appears to be spinning at the correct rate and in the correct direction. When it is rotating faster than 12 Hz but slower than 24 Hz, it appears to be rotating the opposite direction, and at a slower rate. As the rotation rate approaches 24 Hz, the wheel apparently slows down and then stops when the rotation rate equals twice the Nyquist rate. (Note: Because of the symmetry of wagon wheels, this is a slightly simplified picture. Aliasing actually occurs at wheel rotation rates exceeding 12 Hz divided by the number of spokes.)

Power Spectrum

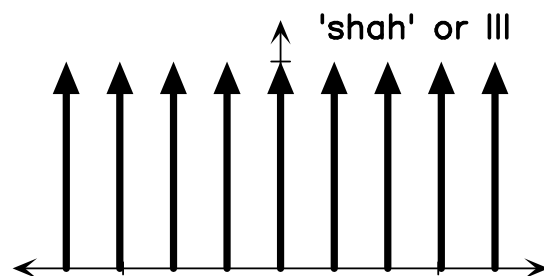
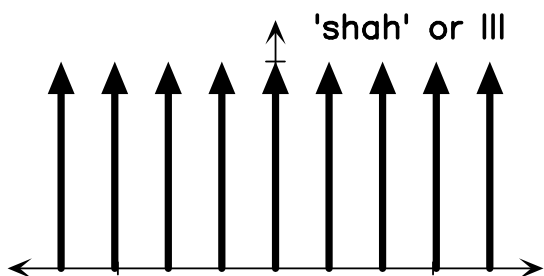
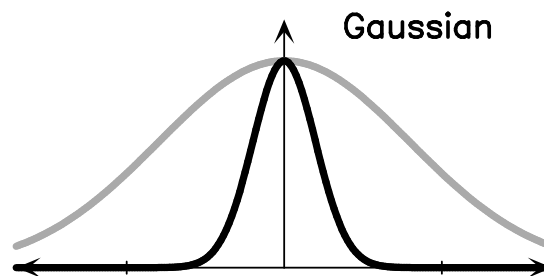
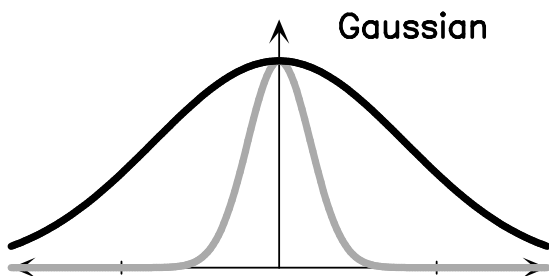
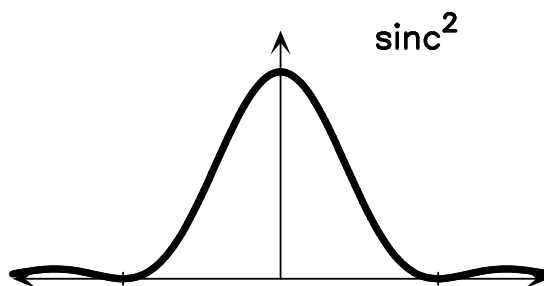
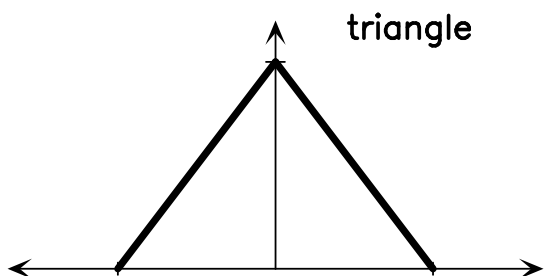
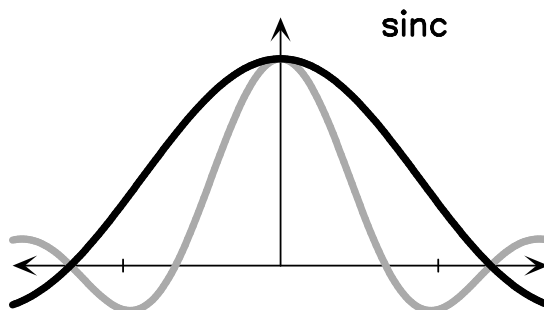
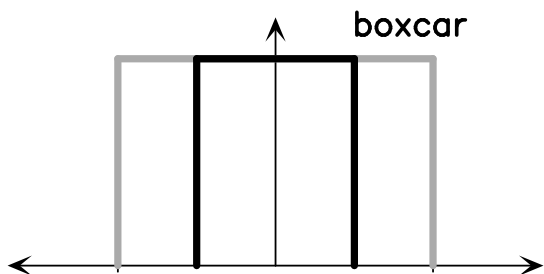
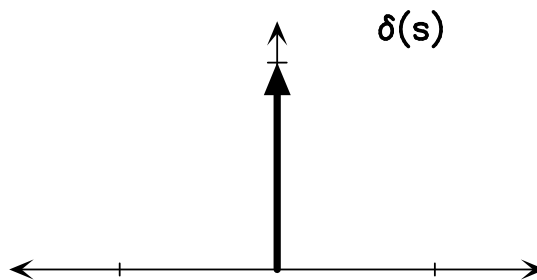
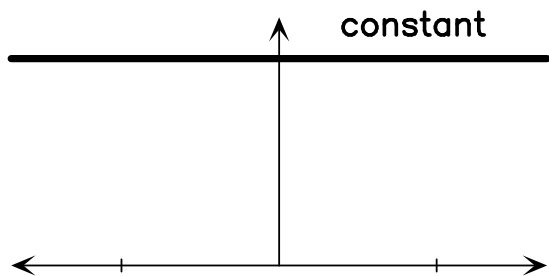
A useful quantity in astronomy is the power spectrum $\overline{F(s)}F(s) = |F(s)|^2$. The power spectrum contains no phase information from the original function.

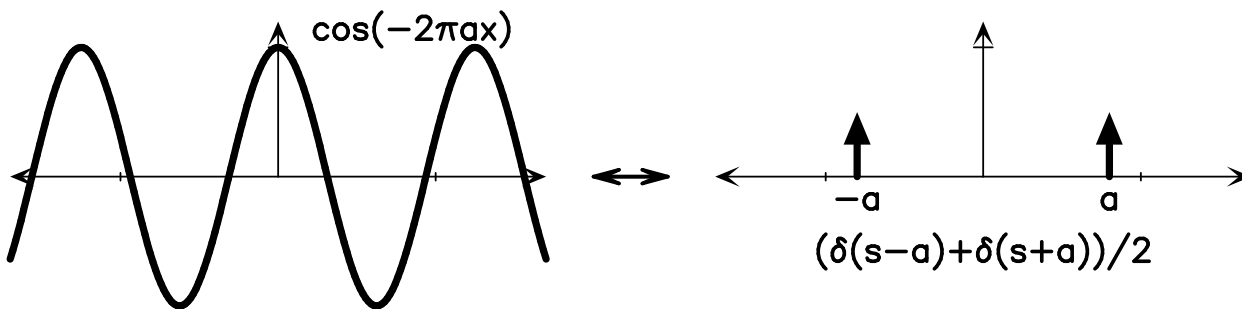
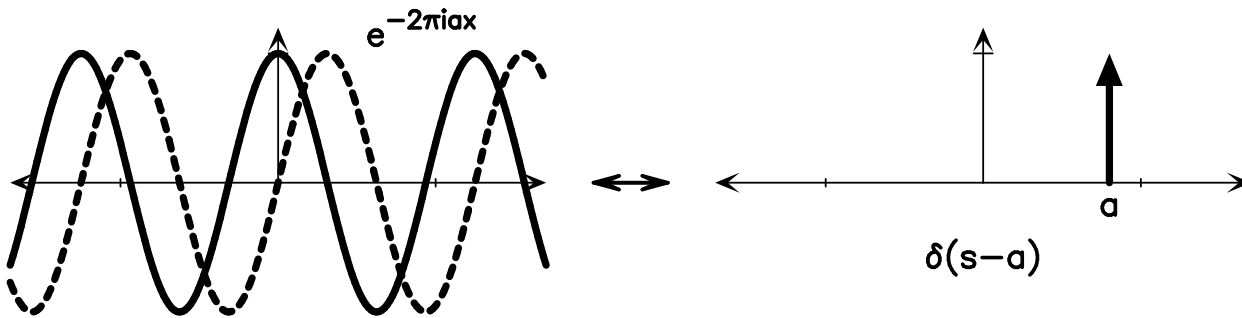
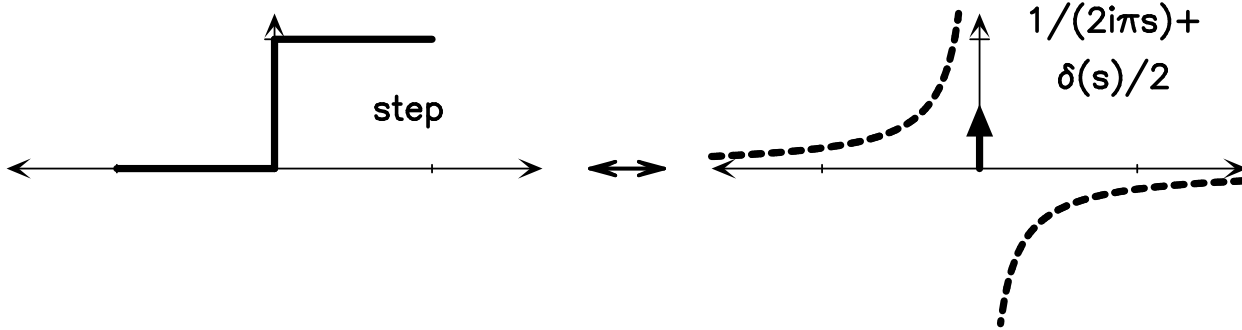
Rayleigh's Theorem (sometimes called Plancherel's Theorem and related to Parseval's Theorem for Fourier series) shows that the integral of the power spectrum equals the integral of the squared modulus of the function (i.e. the energies in the frequency and time domains are equal):

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |F(s)|^2 ds \quad (\text{SF7})$$

Basic Transforms

The following images show basic Fourier transform pairs. These can be combined using the Fourier transform theorems below to generate the Fourier transforms of many different functions. There is a nice little Java applet [here](#) that lets you experiment with various simple DFTs.





Basic Fourier Theorems

Addition Theorem: The Fourier transform of the addition of two functions $f(x)$ and $g(x)$ is the addition of their Fourier transforms $F(s)$ and $G(s)$. This basic theorem results from the linearity of the Fourier transform. A special case of the addition theorem states that if a is a constant, then $af(x) \Leftrightarrow aF(s)$.

$$f(x) + g(x) \Leftrightarrow F(k) + G(k) \quad (\text{SF8})$$

Shift Theorem: A function $f(x)$ shifted along the x-axis by a to become $f(x-a)$ has the Fourier transform $e^{-2\pi ias} F(s)$. The magnitude of the transform is the same, only the phases change.

$$f(x - a) \Leftrightarrow e^{-2\pi i a s} F(s) \quad (\text{SF9})$$

Similarity Theorem: For a function $f(x)$ with a Fourier transform $F(s)$, if the x-axis is scaled by a constant a so that we have $f(ax)$, the Fourier transform becomes $|a|^{-1} F(s/a)$. In other words, a "wide" function in the time-domain is a "narrow" function in the frequency-domain. This is the basis of the uncertainty principle in quantum mechanics and the diffraction limits of radio telescopes.

$$F(ax) \Leftrightarrow \frac{F(s/a)}{|a|} \quad (\text{SF10})$$

Modulation Theorem: The Fourier transform of a function $f(x)$ multiplied by $\cos(2\pi f x)$ is $\frac{1}{2}F(s - f) + \frac{1}{2}F(s + f)$. This theorem is very important in radio astronomy as it describes how signals can be "mixed" to different intermediate frequencies (i.e. IFs).

$$f(x) \cos(2\pi f x) \Leftrightarrow \frac{1}{2}F(s - f) + \frac{1}{2}F(s + f) \quad (\text{SF11})$$

Derivative Theorem: The Fourier transform of the derivative of a function $f(x)$, $f'(x)$, is $i2\pi s F(s)$.

$$f'(x) \Leftrightarrow i2\pi s F(s) \quad (\text{SF12})$$

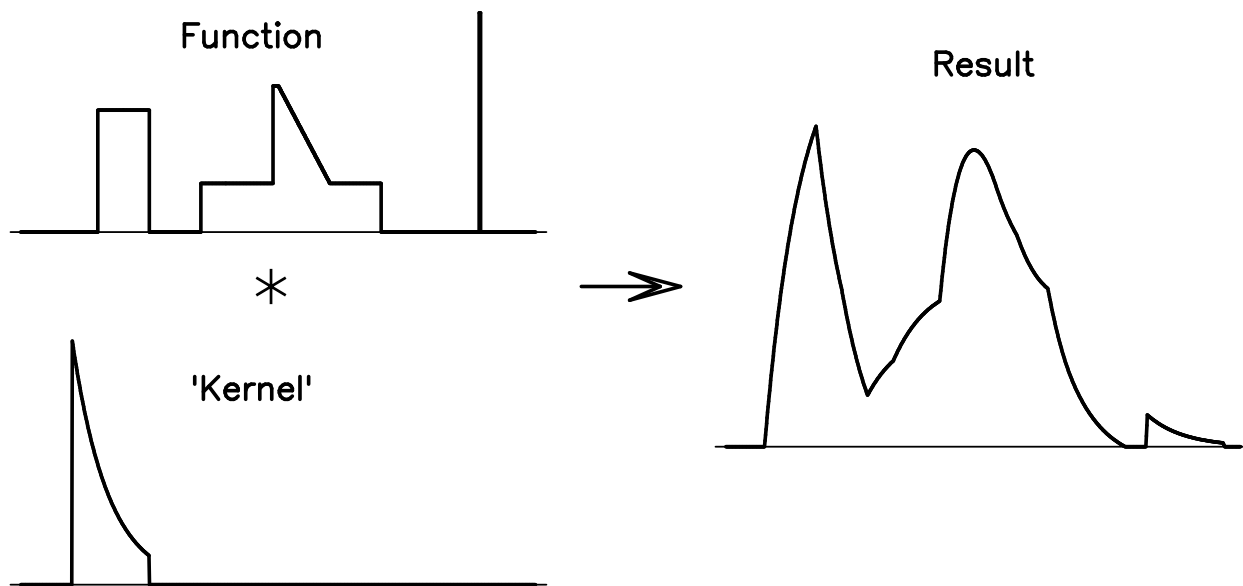
Convolution and Cross-Correlation

Convolution shows up in many aspects of astronomy, most notably in the point-source response of an imaging system and in interpolation. Convolution, which we will represent by $*$ (the symbol \otimes is also frequently used for convolution), multiplies one function f by the time-reversed function g , shifts g by some

amount x , and integrates from $-\infty$ to $+\infty$. The **convolution** $h(x)$ of the functions f and g is a linear functional of $f(x)$ and $g(x)$ defined by:

$$h(x) = f * g \equiv \int_{-\infty}^{\infty} f(u)g(x-u) du \quad (\text{SF13})$$

Convolution Example



Notice how the delta-function like part of the top function produces an image of the convolution kernel in the result. For a time series, that kernel image defines the impulse response of the system. For an antenna or imaging system that would be the point-source response.

A very nice applet showing how convolution works can be found [here](#) (there is also a [discrete one](#)). And here is a [different one](#).

The **convolution theorem** is extremely powerful and states that the Fourier transform of the convolution of two functions is the product of their individual Fourier transforms:

$$f * g \Leftrightarrow F \cdot G \quad (\text{SF14})$$

Cross-correlation is a very similar operation to convolution, except that the "kernel" is not time-reversed during the operation. Cross-correlation is used extensively in interferometry and aperture synthesis imaging, and is also used to perform optimal "matched-filtering" of data to find and identify weak signals. Cross-correlation is represented by the pentagram symbol \star and defined by:

$$f \star g \equiv \int_{-\infty}^{\infty} f(u)g(u-x) du \quad (\text{SF15})$$

In this case, unlike for convolution, $f(x) \star g(x) \neq g(x) \star f(x)$.

Very much related to the convolution theorem, the **cross-correlation theorem** states that the Fourier transform of the cross-correlation of two functions is equal to the product of the individual Fourier transforms, where one of them has been complex conjugated:

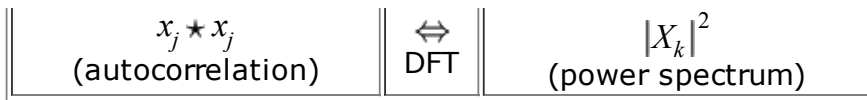
$$f \star g \Leftrightarrow \bar{F} \cdot G \quad (\text{SF16})$$

Autocorrelation is a special case of cross-correlation with $f \star f$. The related autocorrelation theorem is also known as the **Wiener-Khinchin theorem** and states

$$f \star f \Leftrightarrow \bar{F} \cdot F = |F|^2 \quad (\text{SF17})$$

In words, the Fourier transform of an autocorrelation function is the power spectrum, or equivalently, the autocorrelation is the inverse Fourier transform of the power spectrum. Many radio-astronomy instruments compute power spectra using autocorrelations and this theorem. The following table summarizes the relations between a function, its Fourier transform, its autocorrelation, and its power spectrum:

x_j (function)	\Leftrightarrow DFT	X_k (transform)
\Downarrow		\Downarrow



One important thing to remember about convolution and correlation using DFTs is that they are *cyclic* with a period corresponding to the length of the longest component of the convolution or correlation. Unless this periodicity is taken into account (usually by zero-padding one of the input functions), the convolution or correlation will wrap around the ends and possibly "contaminate" the resulting function.

Other links

[Fourier series applet](#)

[Tones, harmonics, filtering, and sounds](#)

[A nice online book on the mathematics of the DFT](#)
