Scaling the Fourier transformation (FT)

Quick summary

Important note on DFT coefficients $\{X_k\}$of real signals:
Per definition, the DFT returns complex coefficients for both positive and negative frequencies, but only positive frequencies can be interpreted physically.
The amplitude of a signal component depends on the sampling frequency Fs, and on the number of data points (N) in the time series.
If N is even, the **amplitude of a signal component** is

$|X_k|$ for X_k at frequency=0 Hz and at frequency=F_S/2 Hz, and
$2\cdot|X_k|$ for X_k at all other positive frequencies.

If N is odd, the DFT_amplitude of a signal component is

$|X_k|$ for X_k at frequency=0 Hz, and
$2\cdot|X_k|$ for X_k at all other positive frequencies

The **unit of a DFT transform** is the same as that of the original signal: [Unit]
[Unit]/Hz may be used, too, but it is mathematically not correct. Some publications show [Unit]/sqrt(Hz). This might come from the units of the energy spectrum (see below) but it isn't justified mathematically and the DFT coefficients are expected to be scaled appropriately. It might also come from the power spectral density (PSD, see below) but then the DFT coefficients must be scaled appropriately here, too, and taking the square root of the PSD is meaningless anyway.

To get the **energy spectrum** correctly:
If N is even, the PS value of a signal component is

$1/(Fs) \cdot |X_k|^2$ for X_k at frequency=0 Hz and at frequency=F_S/2 Hz, and
$1/(Fs) \cdot 2\cdot|X_k|^2$ for X_k at all other positive frequencies.

If N is odd, the PS value of a signal component is

$1/(Fs) \cdot |X_k|^2$ for X_k at frequency=0 Hz, and
$1/(Fs) \cdot 2\cdot|X_k|^2$ for X_k at all other positive frequencies.

The unit is $[Unit]^2$ / Hz if the data is a time series and the sampling interval is given in seconds.

To get the **power spectrum** correctly:
If N is even, the PS value of a signal component is

$1/(N^2) \cdot |X_k|^2$ for X_k at frequency=0 Hz and at frequency=F_S/2 Hz, and
$1/(N^2) \cdot 2\cdot|X_k|^2$ for X_k at all other positive frequencies.

If N is odd, the PS value of a signal component is

$1/(N^2) \cdot |X_k|^2$ for X_k at frequency=0 Hz, and
$1/(N^2) \cdot 2\cdot|X_k|^2$ for X_k at all other positive frequencies.

The unit is simply $[Unit]^2$

To get the **power spectral density** correctly:
If N is even, the PS value of a signal component is

$1/(N\cdot Fs) \cdot |X_k|^2$ for X_k at frequency=0 Hz and at frequency=F_S/2 Hz, and
$1/(N\cdot Fs) \cdot 2\cdot|X_k|^2$ for X_k at all other positive frequencies.

If N is odd, the PS value of a signal component is

$1/(N\cdot Fs) \cdot |X_k|^2$ for X_k at frequency=0 Hz, and
$1/(N\cdot Fs) \cdot 2\cdot|X_k|^2$ for X_k at all other positive frequencies.

The unit is $[\text{Unit}]^2$ / Hz if the data is a time series and the sampling interval is given in seconds.

Details

Analitical FT and inverse FT for continuous functions (according to Wolfram MathWorld)
http://mathworld.wolfram.com/FourierTransform.html

$$f(x) = \int_{-\infty}^{\infty} F(k)\, e^{2\pi i k x}\, dk$$
$$F(k) = \int_{-\infty}^{\infty} f(x)\, e^{-2\pi i k x}\, dx.$$

x == time for functions of time and k == frequency. Note that we can talk about frequency in units of Hz only if x is time and if it is measured in seconds. Otherwise using the term 'cycles per one unit of x' for k is more correct.

f(x) is often a real function but in any case F(k) is a complex function.

At each frequency (k), F(k) can be written in the Euler notation as $F(k) = A(k)\cdot\exp(2\pi\cdot i\cdot\theta(k))$
Here A(k) is the amplitude and θ(k) is the phase of the corresponding sinusoidal component $A(k)\cdot\cos(2\pi\cdot k\cdot x + \theta(k))$, and i is the imaginary unit sqrt(-1).

If values of f(x) are measured in volts (V) then F(k) has the units volt·seconds (V·s) since f(x) is multiplied by the delta time (dx) during integration. V·s is the same as V/Hz.

For a series of values measured (sampled) at discrete points of time (i.e., a time series) $\{f_k = f(x=t_k)\}$, the discrete Fourier transform (DFT) can be calculated.
According to Wolfram MathWorld
(http://mathworld.wolfram.com/DiscreteFourierTransform.html)

discrete Fourier transform $F_n = \mathcal{F}_k\left[\{f_k\}_{k=0}^{N-1}\right](n)$ as

$$F_n \equiv \sum_{k=0}^{N-1} f_k\, e^{-2\pi i n k/N}.$$

The inverse transform $f_k = \mathcal{F}_n^{-1}\left[\{F_n\}_{n=0}^{N-1}\right](k)$ is then

$$f_k = \frac{1}{N}\sum_{n=0}^{N-1} F_n\, e^{2\pi i k n/N}.$$

The result is a series of complex values $\{F_n\}$, i.e., the Fourier series, where n corresponds to different frequencies. Note that in general, DFT works with a data series and does not care about the size of the spacing between the values of f_k (the spacing, nevertheless must be constant). Therefore, the assignment of different values of n to different frequencies (in Hz) is valid only if f_k is a time series and time is measured in seconds.

If the time series consists of N samples and the sampling interval is Δt sec (which corresponds to a sampling frequency $F_S=1/\Delta t$), the DFT will also have N complex values corresponding to N frequencies that have a spacing of F_S/N Hz.
If N is even, the set of frequencies is
$\{ -(N/2 - 2)\cdot F_S/N ; -(N/2 - 3)\cdot F_S/N ; \ldots; -2\cdot F_S/N ; -F_S/N ;$
$0 ; F_S/N ; 2\cdot F_S/N ; \ldots ; (N/2 - 2)\cdot F_S/N ; (N/2 - 1)\cdot F_S/N ; (N/2)\cdot F_S/N \}$
If N is odd, the set of frequencies is
$\{ -((N-1)/2) \cdot F_S/N ; -((N-1)/2 - 1)\cdot F_S/N ; -((N-1)/2 - 2)\cdot F_S/N ; \ldots; -2\cdot F_S/N ; -F_S/N ;$
$0 ; F_S/N ; 2\cdot F_S/N ; \ldots ; ((N-1)/2 - 2)\cdot F_S/N ; ((N-1)/2 - 1)\cdot F_S/N ; ((N-1)/2)\cdot F_S/N \}$
Note that there are also negative frequencies.

There is no integration in DFT formula, only summation in which the values of the time series are multiplied by a dimensionless exponential term. So the unit of the F_n values is practically the same as that of the f_k values. *(To be consistent with the continuous FT, implicit multiplication of each term in the sum by a unit sampling interval can be considered. In this case, the unit of the Fourier series will be unit_of_ f_k × unit_of_sampling_interval. If f_k is measured in volts and time is measured in seconds then frequencies will be in Hz and the unit of F_n values will be V·s or V/Hz which is the same. But this interpretation is mathematically not correct so it is generally not used.)*

Note that in practice, coded functions which calculate the DFT are often called as FFT. Be aware of that FFT is originally an acronym of Fast Fourier transform, which is a tricky implementation of the DFT which works very fast on data series in which the number of elements is an integer power of 2.

Also note that different code implementations of the DFT and inverse DFT can apply the scaling factor 1/N (displayed in the inverse formula above) at different places. Some routines apply it rather at the forward transformation; others may apply it at both conversions using the square root of it. Always be sure what is the case with the actual code you use. The easiest test is to have the DFT of the vector [1,0,0,0] calculated. This is a Kronecker delta function. If the absolute values of the result are [1,1,1,1], then the routine uses the DFT formula from Wolfram Math World shown above.

Further note that the order of F_n values in the output (often referred to as the ‘packing’ of the result) can vary in different code implementations of the DFT. Most frequently, the series of returned F_n values starts with the value that corresponds to zero frequency. Then come the values which correspond to greater and greater frequencies. The next is value belonging to the most negative frequency, and finally come the values which correspond to negative frequencies closer and closer to zero. In any case, it is better to check the documentation of the actual DFT implementation.

If one wants to work with the sinusoidal decomposition of a time series (i.e., with the spectrum), it must be verified that the amplitudes of the sinusoidal components are correct so that the time series can in fact be reconstructed by the superposition of the components with the corresponding amplitudes and phases at all frequencies. An easy test is to create a sinusoid of a given amplitude for an integer number of cycles and then to calculate its DFT. The amplitude at the matching frequency must be accurate. For example:
$y(t) = \text{amplitude}\cdot\cos(2\cdot\pi \cdot\text{frequency}\cdot t + \text{phase})$

amplitude = 1.0
frequency = 10 (Hz)
phase = 0 (rad)
$F_S > 20$ Hz is required in this case because of the Nyquist sampling rule.
For $F_S = 100$ Hz, t = {0, 0.01, 0.02, … , 0.98, 0.99} sec.
The frequencies will be {0, +/-1, +/-2, … , +/-49, 50} Hz.
When the scaling is right, the absolute value of the coefficient corresponding to 10 Hz is 1.0.

One quite probably needed correction to the raw DFT is that amplitudes need to be doubled at frequencies which have a negative pair.

If the inverse DFT is needed after the transformation of the spectrum, one must put extra care to scale the amplitudes back and to pack the real and imaginary parts of the F_n values the right way before the data is sent to the routine of the inverse DFT. (If the back-transformed time series is supposed to consist of real values, amplitudes of the F_n values corresponding to both positive and negative frequencies must be equal and half the true amplitude. Additionally, the phases at negative frequencies must be the same as those at the corresponding positive frequency but multiplied by -1.)

The energy of a continuous signal of finite length is given by

$$E = \int_{-\infty}^{\infty} |x(t)|^2 \, dt.$$

If a series of data is considered, the (momentary) energy is usually taken as the square of the data value at each point of the data series, multiplied by the sampling interval. The total energy in the full interval covered by all of the data points is the sum of the squared values in the data series, multiplied by the sampling interval. The unit for the energy at each point is, therefore, $V^2 \cdot s$ or V^2/Hz (provided the data was measured in volts).

In practice, one always applies DFT on a data series of finite length. DFT handles the original dataset as if it was an infinite periodic data series where one period is the actually examined data series. As an infinite data series, its energy is infinite. On the other hand, its power can be defined as the energy over a unit period.

The power spectrum (PS) of a data series describes the distribution of the total power into frequency components composing that signal (https://en.wikipedia.org/wiki/Spectral_density). The PS can be calculated by using Parseval's theorem for DFT.

Parseval's theorem applied to the FT says that the sum (or integral) of the square of a function is equal to the sum (or integral) of the square of its transform (https://en.wikipedia.org/wiki/Parseval%27s_theorem).
For DFT this statement is described by the following formula

$$\sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2$$

where $X[k]$ is the DFT of $x[n]$, both of length N.

This is easy to understand if the DFT of a Kronecker delta function is considered. The raw DFT of a single unit impulse anywhere in a series of N points (i.e., a Kronecker delta function) has unit amplitude at all (N) frequencies. The linearity of the FT ensures that any scaling of such an impulse is exactly the same both in the original data and in the DFT. The formula of the theorem simply appears if we consider that any data series can be composed as the superposition of a series of scaled Kronecker delta functions.

Parseval's theorem allows calculating the full energy of the data series from the raw coefficients of the Fourier series $\{X_k\}$, and it gives also the distribution of the energy among the frequencies participating in the decomposition. Note that to have the energy, each of the squared coefficients must be multiplied by the sampling time interval Δt or $1/F_S$.

To get the power, the energy over a unit period must be given. The momentary energy $|x_n|^2 \cdot \Delta t = |x_n|^2/F_S$ varies over the time series but the FT components are sinusoids in which time is not defined. Therefore, scaling of the signal must be determined in the original data space (i.e., in the time domain), then applied both in the time domain and in the transformed data (i.e., in the frequency domain). The average energy must be used, which means that the full energy must be divided by the time covered by the data series $N \cdot \Delta t$ or N/F_S. Division by N/F_S results in division by N^2 on the right side of the equation of Parseval's theorem.

This is almost the complete formula for the PS. Note, however, that most frequencies are represented twice in the output of the raw DFT: with a positive and also with a negative sign. Since only positive frequencies have physical meaning, only these should be considered. But then the energy of the negative parts should be added to the positive parts to fulfill the requirement for the conservation of energy.

This means that if the series $\{x_n\}$ is transformed into the raw Fourier series $\{X_k\}$ by DFT, the PS is calculated as $\{ (1/N^2) \cdot |X_k|^2$ for frequency 0 and for frequency $F_S/2$ if N is even; and $(1/N^2) \cdot 2 \cdot |X_k|^2$ for other frequencies $\}$. In this case, the unit of the PS values is V^2, if data values are measured in volts.

Note that the factor 2 appears also if the true amplitude of a sinusoidal component is used to calculate the power. Apart from the division by N, the true amplitude usually obtained by multiplying the amplitude from the raw DFT output by 2, i.e., $2 \cdot |X_k|$. But this component is a sinusoid and as such, its energy can be made equivalent to that of a constant signal. To get the amplitude of the equivalent constant signal, the original amplitude must be divided by sqrt(2). So the equivalent amplitude is now $2 \cdot |X_k|/\mathrm{sqrt}(2) = \mathrm{sqrt}(2) \cdot |X_k|$. The square of this is exactly $2 \cdot |X_k|^2$.

Taking it one step further, the power spectral density (PSD) of the signal describes the power present in the signal as a function of frequency, per unit frequency. The PSD can be calculated from the power spectrum by scaling the values of the PS to the unit frequency, i.e., by dividing the values in the PS by the frequency spacing. If the sampling frequency was F_S, then the frequency spacing is F_S/N (N is the number of data in the time series).

With that, the PSD is calculated as $\{ |X_k|^2/(N \cdot F_S)$ for frequency 0 and for frequency $F_S/2$ if N is even; and $2 \cdot |X_k|^2/(N \cdot F_S)$ for other frequencies $\}$. Note that $\{X_k\}$ are the raw Fourier series coefficients after the DFT. In this case, the unit of the PSD values is V^2/Hz, if data values are measured in volts and the time is measured in seconds.

This is consistent with the description in this paper:
http://edoc.mpg.de/get.epl?fid=55356&did=395068&ver=0
But the paper discusses the extension of the procedure of calculating PS and PSD so that different window functions are used so that the leakage in the resulting spectra is smaller.