# Chapter 1

# A mathematical background: Fourier Theory

"Fourier is a mathematical poem."

*Lord Kelvin*

## 1.1 Analysis of periodic functions

Fourier invented his series method in order to solve the heat-diffusion equation under specified boundary conditions. In this chapter, we shall mainly use the spatial representation, $f(x)$, since later we shall need to extend it to two- and three-dimensional functions. Spatial frequency $\frac{1}{\lambda}$ is simply the inverse of wavelength $\lambda$ of $f(x)$.

### 1.1.1 Orthogonal expansions

We want to start with orthogonal expansions because the Fourier Series is one of them. Fourier Series are used primarily to represent periodic functions. However, we shall see that they can also be used to represent non-periodic functions. So, by choosing an orthogonal set of basis functions $\{\psi_n(x)\}$, we can represent almost any arbitrary $f(x)$ on some interval $(x_1, x_2)$ with the expansion

$$f(x) = c_1\psi_1(x) + c_2\psi_2(x) + c_3\psi_3(x) + ... = \sum_{n=1}^{\infty} c_n\psi_n(x) \tag{1.1}$$

where the coefficients $c_n$ weight each term of the series. The terms may be finite or infinite, but for it to be useful we need to find the correct basis set of functions. Additionally, we need to be able to determine the value of the coefficients without being required to know the value of any other coefficient. Luckily for us, these properties can be realized as long as the basis functions $\psi_n(x)$ form a *complete orthogonal set.*

We will focus on the basis functions that will also generate the space where the Fourier Series live. One of the most important set of basis functions for the Fourier series is the complex exponential form in Eq.1.2.

$$\psi_n(x) = e^{i2\pi n\nu_0 x}, \quad n = 0, \pm 1, \pm 2, ... \tag{1.2}$$

They are particularly important because they are eigenfunctions of linear shift-invariant (LSI) operators, which are the operators used to describe LSI systems denoted by Ł{ }.

Now, we know that the eigenfunctions of a linear shift-invariant system are passed by the system unchanged in form, but possibly changed in magnitude and position (Davis). What this means is that in LSI optical systems the only effect caused by a shift in the position of the input is an equal shift in the position of the output, but the magnitude may be changed. The optical systems we will be describing are LSI.(Gaskill)

In addition to the complex exponentials above, the following functions in Eq.1.3 form complete orthogonal sets over the same intervals, and they too are eigenfunctions.

$$\begin{aligned} \psi_n(x) &= \sin(2\pi n\nu_0 x), \quad n = 0, 1, 2, ... \\ \psi_n(x) &= \cos(2\pi n\nu_0 x), \quad n = 0, 1, 2, ... \end{aligned} \tag{1.3}$$

Evidently, these functions are related to the complex exponentials by Euler's formula for when a complex number is given in polar form $u = re^{i\phi}$ and we need to determine its components. $u = r(cos(\phi) + isin(\phi))$

It should be specified that the orthogonality of both sets is limited any interval equal to an integral number of periods of the first-order terms,

$$x_1 - x_2 = k\frac{1}{\nu_0}, k = 1, 2, 3...$$

We know that when the input of a LSI system is an eigenfunction of the system, the output is simply the product of the input and a complex constant of proportionality $H(\xi_0)$., where $\xi$ is an arbitrary real constant.

Let's consider a "special" LSI system with a real-valued input and output. For this system a cosine input yields a cosine output, possibly attenuated and shifted. the complete proof can be found in (Davis)

But what happens when we want to find the output of an arbitrary input signal? It is apparent that we can deconstruct this input signal into its Fourier components due to the fact that they are complex exponentials of the form of Eq.1.2. We know that these exponentials are eigenfunctions of LSI systems. So, if we also know the transfer function we can determine how much each Fourier component is attenuated and phase shifted in passing through the system. From there we apply the superposition principle and add all of the Fourier components to obtain the overall response. The details are more intricate that what was shown and can be consulted in references (A. Lipson, S. G. Lipson, and H. Lipson) and (Goodman).

Now that we understand the algebraic context of Fourier Theory let's continue with Fourier's Theorem.

### 1.1.2 Fourier's theorem

Fourier's theorem states that any periodic function $f(x)$ can be expressed as the sum of a series of sinusoidal functions which have wavelengths that are integral fractions of the wavelength of $f(x)$. To make this statement complete, zero is counted as an integer, giving a constant leading term to the series:

$$f(x) = \frac{1}{2}C_0 \quad + \sum_{n=1}^{\infty} C_n \cos(nk_0 x + \alpha_n) \tag{1.4}$$

Where $k_0 \equiv 2\pi/\lambda$ is the *fundamental spatial frequency* and the $n$s are called the orders of the terms, which are harmonics. However, the functions are not always identical when the number of terms becomes infinite; there are examples that do not converge to the required function, but the regions of error must become vanishingly small (A. Lipson, S. G. Lipson, and H. Lipson)

### 1.1.3 Fourier Series

Each term in the series Eq.1.4 has two Fourier coefficients, an amplitude $C_n$ and a phase angle $\alpha_n$. The latter quantity provides the degree of freedom necessary for relative displacements of the terms of the series along the x-axis. The determination of these quantities for each term of the series is called *Fourier analysis*. Another way to express the Fourier coefficients is to write Eq.1.4 as a sum of sine and cosine terms:

$$f(x) = \frac{1}{2}A_0 \quad + \sum_{1}^{\infty} A_n \cos(nk_0 x) + \sum_{1}^{\infty} B_n \sin(nk_0 x) \tag{1.5}$$

where $A_n = C_n cos(\alpha_n), B_n = C_n sin(\alpha_n)$.

The Fourier series is written in complex notation as,

$$f(x) = \sum_{1}^{\infty} F_n e^{(nk_0 x)} \tag{1.6}$$

where $F_0 = \frac{1}{2}A_0$. We have assumed so far thet the function $f(x)$ is real. Nonetheless, a complex function $f(x)$ can be represented by complex coefficients $A_n$ and $B_n$.

## 1.2 Analysis of non-periodic functions

With everything we've learned we might ask why Fourier methods are of any importance, since they apply to periodic functions only. The answer is that the theory has an extension, not visualized by Fourier himself, to non-periodic functions. The extension is based upon the concept of the *Fourier transform*.

### 1.2.1 The Fourier Transform (FT)

When our interest turns to non-periodic functions we construct a wave of wavelength $\lambda$ in which each unit consists of some non-periodic function.

We can always make $\lambda$ so large that an insignificant amount of the function lies outside the one-wavelength unit. Now allow $\lambda$ to increase without limit, so that the repeats of the non-periodic function separate further and further.

What happens to the function $F(k)$? The spikes approach one another as $\lambda$ increases, but one finds that the envelope of the tips of the spikes remains invariant; it is determined only by the unit, the original non-periodic function. In the limit of $\lambda \to \infty$ the spikes are infinitely close to one another, and the function F(k) has just become the envelope. This envelope is called the Fourier transform of the non-periodic function.

Admittedly, this suggests that the Fourier series for a non-periodic function is a set of spikes at discrete but infinitesimally spaced frequencies rather than a continuous function. We don't know if in the limit $\lambda \to \infty$ the function becomes continuous; although, that's a detail physicists can overlook.

We now define the Fourier transform of a function $f(x)$ as

$$\mathscr{F}(k) = \int_{-\infty}^{\infty} f(x) e^{-ikx} dx \tag{1.7}$$

which is a continuous function of the spatial frequency k. Notice that the $1/2\pi$ has been dropped; this has no physical significance.

**Shift Property of the FT**

If $f(x) \underset{FT}{\longrightarrow} \mathscr{F}(k)$ and $x_0$ is a real constant, even zero, then.

$$\begin{aligned} \mathscr{F}\{f(x - x_0)\} &= \int_{-\infty}^{\infty} f(\alpha - x_0) e^{-j2\pi\xi\alpha} d\alpha \\ &= \int_{-\infty}^{\infty} f(\beta) e^{-j2\pi\xi(\beta + x_0)} d\beta \\ &= e^{-j2\pi x_0 \xi} \int_{-\infty}^{\infty} f(\beta)\, e^{-j2\pi\xi\beta} d\beta \\ &= e^{-j2\pi x_0 \xi} F(\xi) \end{aligned} \tag{1.8}$$

Thus, the FT of a shifted function is the transform of the unshifted function multiplied by an exponential factor having linear phase.This term causes each Fourier component to be shifted in phase by an amount proportional to the product of its frequency and the shift distance $x_0$.

### 1.2.2 Nyquist–Shannon sampling theorem

Also known by the names Nyquist–Shannon–Kotelnikov, Whittaker- Shannon–Kotelnikov, Whittaker–Nyquist–Kotelnikov–Shannon, and cardinal theorem of interpolation, The Sampling Theorem is the bridge between continuous-time signals and discrete-time signals.

Sampling is a process of converting a signal (for example, a function of continuous time and/or space) into a numeric sequence (a function of discrete time and/or space). This introduces the concept of a **sample rate** that is sufficient for perfect fidelity. In other words, it establishes the sufficient sample rate for a discrete sequence of samples to be able to capture all the information from a continuous-time signal of finite bandwidth. Meaning, that no actual information is lost in the sampling process.

It can only be applied to bandlimited functions to a given bandwidth B, in the low pass sense. They are mathematical functions that have Fourier Tranforms equal to zero outside of a finite region of frequencies,

if there is a $B > 0$ such that,

$$\mathscr{F}(k) = \mathscr{F}(k) \prod \left( \frac{k}{2B} \right)$$

it then follows that,

$$f(x) = \int_{-B}^{B} \mathscr{F}(k) e^{j2\pi kt} dk$$

The signal can be expressed by a cardinal series in most of the cases,

$$f(x) = \sum_{n=-\infty}^{\infty} f\left(\tfrac{n}{2B}\right) sin(2\pi Bx - n) \tag{1.9}$$

The ability to express a continuous signal in terms of its samples is the fundamental statement of the sampling theorem,

If a function f(x) contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced 1/(2B) seconds apart.

A sufficient sample-rate is therefore 2B samples/second, or anything larger. Equivalently, for a given sample rate $f_s$, perfect reconstruction is guaranteed possible for a bandlimit $B < f_s/2$.

An easy proof can be given that directly relates the Sampling Theorem to Fourier Theory.

Since

$$\mathscr{F}(k) = 0 \quad when \quad |k| > B,$$

it can be replicated to form a periodic function in te frequency domain with period $2B$. This periodic function can be expressed as a Fourier Series. The result of the series for $k > B$ is $\mathscr{F}(k)$. Using the results in Secc. 1.1 we know

$$\mathscr{F}(k) = \textstyle\sum_{n=-\infty}^{\infty} C_n e^{\frac{-j\pi nk}{B}} \prod \left(\frac{n}{2B}\right) \tag{1.10}$$

Where the Fourier coefficients are

$$C_n = \frac{1}{2B}\int_{-B}^{B} \mathscr{F}(k) e^{\frac{j\pi nk}{B}} dk = \frac{1}{2B} f\left(\frac{n}{2B}\right) \tag{1.11}$$

Substituting into Eq.1.9 and using the inverse FT gives the sampling series in Eq. 1.8, which, as we see here, is the FT dual of the Fourier series. For a more detailed proof please refer to (Bracewell) and (Gallagher).

### 1.2.3 The 2-dimensional Fourier Transform (2FT)

All that has been said so far about Fourier transforms and series in one dimension also applies to two dimensions. In particular, two-dimensional functions (screens) are very important in optics. The transform is defined in terms of two spatial frequency components, $kx$ and $ky$, by a double integral:

$$F(k_x, k_y) = \iint_{-\infty}^{\infty} f(x, y)\, e^{-i(xk_x + yk_y)} dxdy \tag{1.12}$$

### 1.2.4 Fourier algorithms: fast and discrete Fourier transforms

In general, when $f(x)$ is not a simple analytical function, the Fourier transform has to be evaluated numerically. In order to do this, the function $f(x)$ must be defined in a given finite region of size Na, in which it is sampled at N discrete points xn = na. The Fourier transform is then evaluated in a finite regime in reciprocal space, at say M discrete points. In this way, the function is essentially considered as one period of a periodic function, for which the Fourier coefficients are then calculated in the form of a sum for each of the latter values:

$$F(m) = \frac{1}{2\pi} \sum_{n=1}^{N} f(na) e^{\left(-\frac{2\pi i n m}{NM}\right)} \tag{1.13}$$

When N and M are large, this is a very time-consuming calculation to carry out directly. A very efficient algorithm, the fast Fourier transform (FFT), which uses matrix factorization methods to simplify the calculation when $M = N$ is an integer power of 2, was proposed by Cooley and Tukey in 1965 (Brigham (1988)) and is now very widely used. If M and N are not powers of 2, the algorithm pads the regions to make them so. However, the fast Fourier transform is not always the ideal tool. There are many cases where, for example, one needs the Fourier transform in a limited region only, or $M = N$, or the sampling is not uniform; in these cases, the direct evaluation of Eq.1.13 might be more efficient. This is called a Discrete Fourier Transform (DFT).

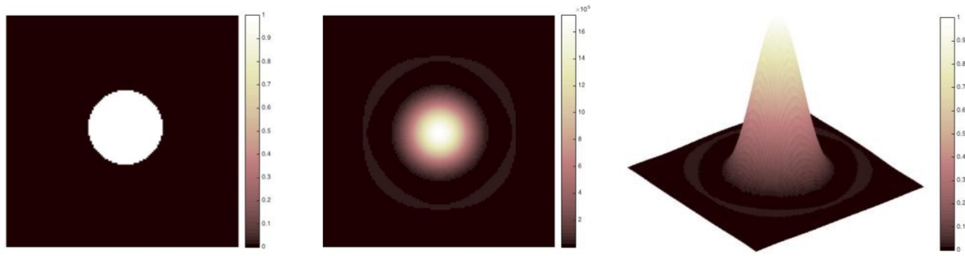

Figure 1.1: Example of FFT2 in Matlab.