TRANSFORMS, ALGORITHMS AND APPLICATIONS

M. J. Soares
Universidade do Minho, Campus de Gualtar 4710-154 Braga, Portugal
jsoares@math.uminho.pt

Abstract - Fourier transforms and other related transforms are an essential tool in applications of science, engineering and technology. In fact, much of the work currently being done in mathematics, physics and engineering has its roots in Fourier's pioneering idea of representing an arbitrary function as the sum of a trigonometric series. The main purpose of these notes is to give a brief overview of some Fourier-related transforms, namely: continuous Fourier transform, Fourier series, discrete Fourier transform, fast Fourier transform (FFT), sine and cosine transforms, Z-transform, Laplace transform, windowed Fourier transform, continuous and discrete wavelet transforms. Our aim is simply to present a summary of these transforms and to describe their main properties and possible applications, and so most of the results are presented with no proof. References containing the proofs and other details about the transforms are always indicated.

Keywords - Fourier transforms, Fourier series, FFT, wavelet transforms.

I. Notations

We start by introducing the main notations that will be used throughout these notes.

• If X is a measurable subset of the real line $\mathbb{R}$, in particular the whole of $\mathbb{R}$, we denote by $L^p(X)$ $(0 < p < \infty)$, the Banach space of the (equivalence classes of) measurable functions f defined in X such that

$$\|f\|_p := \left(\int_X |f(t)|^p dt\right)^{1/p} < \infty. \tag{1}$$

When $p = 2$, this is a Hilbert space with respect to the inner product

$$\langle f, g\rangle := \int_X f(t)\overline{g(t)}dt. \tag{2}$$

(Here and throughout, $\overline{u}$ denotes the complex conjugate of u.)

• When X is a finite interval $X = [a, a + \Omega]$ of length Ω, $\Omega > 0$, we can identify the above space with the space of functions which are periodic of period Ω, i.e. satisfy $f(t + k\Omega) = f(t)$, for all $k \in \mathbb{Z}$ and for almost all t, and are such that $\int_a^{a+\Omega} |f(t)|^p dt < \infty$. In fact, any Ω-periodic function is totally determined by its behaviour on any interval of length Ω and, reciprocally, any function which is only defined on an interval of length Ω can always be periodically extended (with period Ω) to the whole line. We can also think of this space as a space of functions defined on the Ω-torus $\mathbb{T}_\Omega = \mathbb{R}/\Omega\mathbb{Z}$; see Section IV if you are unfamiliar with this type of notation. In this case, it is more convenient to normalize the inner product (2) as

$$\langle f, g\rangle = \frac{1}{\Omega}\int_a^{a+\Omega} f(t)\overline{g(t)}dt. \tag{3}$$

The norm $\|.\|_p$ will also be redefined as

$$\|f\|_p := \left(\frac{1}{\Omega}\int_a^{a+\Omega} |f(t)|^p dt\right)^{1/p}. \tag{4}$$

In order to simplify the notation, we will always write $\int_{\mathbb{T}_\Omega}$ to designate $\frac{1}{\Omega}\int_a^{a+\Omega}$. This means, for example, that the inner product (3) will be written simply as

$$\langle f, g\rangle = \int_{\mathbb{T}_\Omega} f(t)\overline{g(t)}dt. \tag{5}$$

• When X is the discrete set $\mathbb{Z}$, the functions defined on X will simply be two-sided sequences, and we use for them a notation of the type $f = (f[k])_{k\in\mathbb{Z}}$, following the tradition of signal processing literature of using square brackets around a discrete variable. In this case, the integrals in (1) and (2) should be understood with respect to the *discrete measure*, i.e. the norm and inner product are defined, respectively, by

$$\|f\|_p := \left(\sum_{k\in\mathbb{Z}} |f[k]|^p\right)^{1/p} \tag{6}$$

and

$$\langle f, g\rangle := \sum_{k\in\mathbb{Z}} f[k]\overline{g[k]}. \tag{7}$$

These spaces are referred to as the spaces of *p-summable sequences* and denoted by $\ell^p(\mathbb{Z})$.

• Finally, when the set X is discrete and finite, e.g. $X = \{0, 1, \ldots, N-1\}$, the functions on X, which are simply vectors $f = f([k])_{k=0}^{N-1}$, can also be "viewed" as N-periodic sequences on $\ell^p(\mathbb{Z})$ (any p) if we define, for $k \in \mathbb{Z}$, $f[k] = f[k \bmod N]$, where $k \bmod N$ denotes the remainder of the division of k by the integer N. This space can be identified with the space $\ell(\mathbb{Z}_N)$ with $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$; more details, again, in Section IV. Here, naturally, the inner product and norm are the usual Euclidean inner product and norm of vectors in $\mathbb{C}^N$, i.e. they are, respectively

$$\langle f, g\rangle := \sum_{k=0}^{N-1} f[k]\overline{g[k]} \tag{8}$$

and

$$\|f\| := \{\sum_{k=0}^{N-1} |f[k]|^2\}^{1/2}. \tag{9}$$

• A family $\{e_k : k \in \mathbb{Z}\}$ of elements in a Hilbert space H (with inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\| \cdot \|$) is said to be an *orthogonal basis* of H if it satisfies:

1. $\langle e_i, e_j \rangle = 0, \quad i \neq j$;
2. for any $x \in H$, there is a unique sequence of scalars $\hat{x}[k]$ such that

$$\lim_{N\to\infty} \|x - \sum_{k=-N}^{N} \widehat{x}[k] e_k\| = 0. \tag{10}$$

The orthogonality condition implies that the coefficients $\hat{x}[k]$ are necessarily given by

$$\hat{x}[k] = \frac{\langle x, e_k \rangle}{\|e_k\|^2},$$

and we will write (10) simply as

$$x = \sum_{k\in\mathbb{Z}} \frac{\langle x, e_k \rangle}{\|e_k\|^2} e_k. \tag{11}$$

When each vector e_k has unit norm, the basis is said to be *orthonormal* (o.n.). In this case, *Plancherel formula*, which sates an energy conservation, holds:

$$\|x\|^2 = \sum_{k\in\mathbb{Z}} | \langle x, e_k \rangle |^2. \tag{12}$$

We will frequently refer to *Fourier transform* to designate several different mathematical transformations, depending on the nature of the spaces on which they are defined (in other words, depending on the type of *signals* on which they are acting). When necessary, we will be more specific and use terms like *continuous time Fourier transform*, *continuous time Fourier series*, etc. A small table summarizing the Fourier transforms for various settings is given below.

Table of Fourier Transforms

Name	**Domain** of f	**Transform** $F = \hat{f}$ (and Inverse)	**Domain** of $\hat{f}$
CTFT	$\mathbb{R}$	$\hat{f}(\omega) = \int_{\mathbb{R}} f(t)e^{-2\pi i\omega t}dt$ $f(t) = \int_{\mathbb{R}} \hat{f}(\omega)e^{2\pi it\omega}d\omega$	$\mathbb{R}$
CTFS	$\mathbb{T}_\Omega$	$\hat{f}[k] = \int_{\mathbb{T}_\Omega} f(t)e^{-2\pi ikt/\Omega}dt$ $f(t) = \sum_{k\in\mathbb{Z}} \hat{f}[k]e^{2\pi ikt/\Omega}$	$\mathbb{Z}$
DTFT	$\mathbb{Z}$	$\hat{f}(\omega) = \sum_{k\in\mathbb{Z}} f[k]e^{-2\pi ik\omega/\Omega}$ $f[k] = \int_{\mathbb{T}_\Omega} \hat{f}(\omega)e^{2\pi ikt\omega/\Omega}d\omega$	$\mathbb{T}_\Omega$
DTFS	$\mathbb{Z}_N$	$\hat{f}[n] = \sum_{k=0}^{N-1} f[k]e^{-2\pi ikn/N}$ $f[k] = \frac{1}{N}\sum_{n=0}^{N-1} \hat{f}[n]e^{2\pi ikn/N}$	$\mathbb{Z}_N$

CT-continuous time; DT-discrete time; FT-Fourier transform; FS-Fourier series

In Section IV, we wil give a more unified view of these different transforms, briefly describing how they all fit in the more general framework of *Fourier transforms on groups.* For the moment, we will study in more detail each of the above transforms, discussing, in particular the conditions (and the different interpretations) for the *inverse formulas* to hold.

II. Continuous Time Fourier Transform (CTFT)

A. *Fourier transform in* $L^1(\mathbb{R})$

We start by defining the Fourier transform of functions in the space $L^1(\mathbb{R})$.

The Fourier transform (also called *continuous-time Fourier transform* or *integral Fourier transform*) of a function $f \in L^1(\mathbb{R})$ is the function $\hat{f}$ defined by

$$\hat{f}(\omega) := \int_{\mathbb{R}} f(t)e^{-2\pi i\omega t}dt, \quad \omega \in \mathbb{R}. \tag{13}$$

For simplicity, to indicate the correspondence between a function f and its Fourier transform, we use the notation $f \longrightarrow F$.

We consider the following three operators, defined for $a \in \mathbb{R}$:

Translation: $T_a f(t) = f(t-a)$

Modulation: $E_a f(t) = e^{2\pi iat} f(t)$

Dilation: $D_a f(t) = |a|^{-1/2} f(t/a), \ (a \neq 0)$.

The main algebraic and analytic properties of the Fourier transform are summarized in the following two theorems; the proofs can be seen, e.g. in [6].

Theorem 1

1. Linearity $c_1 f_1 + c_2 f_2 \longrightarrow c_1 F_1 + c_2 F_2$.
2. Conjugation $\overline{f}(t) \longrightarrow \overline{F}(-\omega)$.
3. Time shifting $T_a f \longrightarrow E_{-a} F$.
4. Modulation $E_a f \longrightarrow T_a F$.
5. Time dilation $D_a f \longrightarrow D_{1/a} F$.

Theorem 2

Let $f \in L^1(\mathbb{R})$ and let F be its Fourier transform. Then, we have

1. Boundedness *For each $\omega \in \mathbb{R}$, $|F(\omega)| \leq \|f\|_1$.*
2. Continuity *F is (uniformly) continuous on $\mathbb{R}$.*
3. Riemann-Lebesgue Lemma $\lim_{|\omega| \to \infty} F(\omega) = 0$.
4. Time differentiation *Let $f \in C^m(\mathbb{R}) \cap L^1(\mathbb{R})$ be such that $f^{(k)}; k = 1 \ldots, m$, are in $L^1(\mathbb{R})$. Then*
$$f^{(k)}(t) \longrightarrow (2\pi i\omega)^k F(\omega).$$
5. Frequency differentiation *Suppose that $t^m f(t) \in L^1(\mathbb{R})$. Then, $F^{(k)}; k = 1, \ldots, m$, exist and*
$$(-2\pi it)^k f(t) \longrightarrow F^{(k)}(\omega).$$

Another important property of Fourier transform is its behaviour with respect to *convolution*. Recall that the *convolution* $f * g$ of two functions f and g is the function defined by

$$f * g(t) = \int_{\mathbb{R}} f(u)g(t-u)du. \tag{14}$$

We then have the following result:

Theorem 3 (Convolution) *If $f, g \in L^1(\mathbb{R})$, then $f * g \in L^1(\mathbb{R})$ and*

$$f * g \longrightarrow FG.$$

B. *Inversion*

Given a function $g \in L^1(\mathbb{R})$, we define its *inverse Fourier transform* $\check{g}$ by

$$\check{g}(t) := \int_{\mathbb{R}} g(\omega) e^{2\pi i \omega t} d\omega, \quad t \in \mathbb{R},$$

i.e. $\check{g}(t)$ is simply $\hat{g}(-t)$. The name *inverse* Fourier transform is justified by the following theorem, which shows that the function f can be recovered from its Fourier transform, by applying to it the inverse Fourier transform.

Theorem 4 *Let $f \in L^1(\mathbb{R})$ and let $\hat{f}$ denote its Fourier transform. If $\hat{f} \in L^1(\mathbb{R})$, then f is continuous and $f = \check{\hat{f}}$, i.e.*

$$f(t) = \int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i \omega t} d\omega. \tag{15}$$

Note: This theorem establishes a pointwise inversion formula for the Fourier transform under the assumption that $\hat{f} \in L^1(\mathbb{R})$. It should be interpreted in the following sense: the integral on the r.h.s. is defined for every $t \in \mathbb{R}$ and defines a continuous function which coincides with f almost everywhere (a.e.); the pointwise equality is valid for the continuous representative of f.

C. *Fourier transform in $L^2(\mathbb{R})$*

The formula (13) as it stands can not be applied directly to functions in the space $L^2(\mathbb{R})$ (if they are not in $L^1(\mathbb{R})$), so the definition of the Fourier transform for functions in this space (the important space of signals of finite energy) has to be suitably adapted.

The following result is essential for establishing a natural definition for the Fourier transform in $L^2(\mathbb{R})$.

Theorem 5 (Plancherel-Parseval) *If $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then*

$$\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle \quad (\textit{Parseval identity}). \tag{16}$$

In particular, we have

$$\|f\| = \|\hat{f}\| \quad (\textit{Plancherel formula}). \tag{17}$$

The extension of the Fourier transform to $L^2(\mathbb{R})$ is based on the use of the above formulae and the fact that $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is *dense* in $L^2(\mathbb{R})$. This means that, given a function $f \in L^2(\mathbb{R})$, there is a sequence of functions $(f_n)_{n\in\mathbb{N}}$ in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ converging to f (with convergence taken with respect to the norm in $L^2(\mathbb{R})$). This implies that $\|f_m - f_n\|_2 \to 0$ when $m, n \to \infty$. By the linearity of the Fourier transform and the Plancherel formula, we immediately conclude that the sequence $(\widehat{f_n})_{n\in\mathbb{Z}}$ converges to a certain function in $L^2(\mathbb{R})$. This limit function will be called the Fourier transform of f (sometimes called the Plancherel transform) and will also be denoted, as before, by $\hat{f}$ or F. It can be shown that the limit function F does not depend on the choice of the sequence f_n converging to f and, naturally, that it coincides with the usual Fourier transform of f when $f \in L^1(\mathbb{R})$. A standard way of selecting the sequence f_n is to take $f_n = f\,\mathbf{1}_{[-n,n]}$, where $\mathbf{1}_{[a,b]}$ denotes the characteristic function of the interval $[a, b]$, i.e.

$$\mathbf{1}_{[a,b]}(t) = \begin{cases} 1, & t \in [a, b] \\ 0, & \text{otherwise.} \end{cases}$$

If we write $\text{l.i.m.} g_n(t) = g(t)$ to indicate that $\|g_n - g\|_2 \to 0$ when $n \to \infty$, we can thus write, for $f \in L^2(\mathbb{R})$,

$$\hat{f}(\omega) := \text{l.i.m.} \int_{-n}^{n} f(t)e^{-2\pi i\omega t}dt. \tag{18}$$

Note: With a convenient abuse of notation we will still write, when $f \in L^2(\mathbb{R})$,

$$\hat{f}(\omega) = \int_{\mathbb{R}} f(t)e^{-2\pi i\omega t}dt,$$

with the understanding that this is a limiting process as defined above.

It is important to observe that the main properties stated for the Fourier transform of functions in $L^1(\mathbb{R})$ also hold for this extension to $L^2(\mathbb{R})$. The extension of the definition of the inverse Fourier transform $\check{g}$ to functions $g \in L^2(\mathbb{R})$ is, naturally, done in manner analogous to the process described for the Fourier transform, and we also have an inversion theorem for this case.

Theorem 6 (Inversion in $L^2(\mathbb{R})$) *The Fourier transform is a bijective linear operator from $L^2(\mathbb{R})$ into $L^2(\mathbb{R})$. Given $f \in L^2(\mathbb{R})$, we have*

$$f = \check{\hat{f}}.$$

The definition of the Fourier transform can also be extended to a wider class of "objects", the so-called *tempered distributions*; as an example of a tempered distribution we have the Dirac-delta δ. This is a linear functional which acts on a (sufficiently well-behaved function) f by giving its value at zero, i.e.

$$\delta(f) := f(0).$$

The Fourier transform of a tempered distribution is another tempered distribution. In the case of the Dirac-delta, the Fourier transform can be identified with the constant function 1, i.e

$$\hat{\delta} = 1.$$

For more details on Fourier transforms of tempered distributions, see, e.g. [26] or [6].

III. Continuous Time Fourier Series (CTFS)

We now consider the case where the function f to be transformed is in $L^2(\mathbb{T}_\Omega)$, where $\mathbb{T}_\Omega = \mathbb{R}/\Omega\mathbb{Z}$ is the Ω-torus ($\Omega > 0$). It can be shown that the set of functions

$$\gamma_k(t) := e^{2\pi ikt/\Omega}, \quad k \in \mathbb{Z}, \tag{19}$$

is an orthonormal basis of $L^2(\mathbb{T}_\Omega)$ (with respect to the inner product defined by (5)). This means that every function $f \in L^2(\mathbb{T}_\Omega)$ can be written as

$$f(t) = \sum_{k\in\mathbb{Z}} \hat{f}[k] e^{2\pi ikt/\Omega}, \tag{20}$$

where the coefficients $\hat{f}[k]$ are given by

$$\begin{aligned} \hat{f}[k] =& \langle f, \gamma_k \rangle \\ =& \int_{\mathbb{T}_\Omega} f(t) e^{-2\pi ikt/\Omega} dt. \end{aligned} \tag{21}$$

The coefficients $\hat{f}[k],\ k \in \mathbb{Z}$ given by (21), are called the *Fourier coefficients* of the function f and the series on the r.h.s. of (20) is the called the *Fourier series of f.*

The equality (20) is to be interpreted as (cf. 10)

$$\lim_{N\to\infty} \int_0^\Omega |f(t) - \sum_{k=-N}^{N} \hat{f}[k] e^{2\pi ikt/\Omega}|^2 dt = 0$$

and does not necessarily mean that, for every $t \in \mathbb{R}$, the series on the r.h.s. of (20) converges to the value $f(t)$. The problems associated with the pointwise (and uniform) convergence of Fourier series, namely the discussion of the minimum conditions which ensure this type of convergence, have attracted the attention of mathematicians for more than two centuries and had a profound impact on the evolution of the foundations of Analysis; an accessible reference on this subject, with an interesting historical perspective, is [25].

The equality (20) is also known to hold for almost all t; moreover, if the function f is sufficiently well-behaved (e.g. piecewise smooth) then the series converges, at every point t, to the average value

$$\frac{f(t^+) + f(t^-)}{2}.$$

The Fourier series has a typical behaviour near the points of discontinuity; its partial sums overshoot and undershoot the true values $f(t^+)$ and $f(t^-)$, respectively, by about 9% of the total jump $f(t^+) - f(t^-)$. This is the famous *Gibbs phenomenon*, and was observed by Gibbs, for a particular function, in a letter to *Nature* (vol.59, p.606), in 1899.

Note: In fact, this phenomenon had already been described by H. Wilbraham, 51 years earlier [50], although Gibbs was not aware of this. In 1906, M. Bôcher, an American mathematician, proved that this behaviour is a general property of Fourier series in the vicinity of a jump discontinuity; [7].

The computation of the sequence of the Fourier coefficients $\hat{f}[k]$ in the case where f is a periodic function can be seen as the analogue of the computation, for a function f with no periodicity, of $\hat{f}(\omega)$, for all $\omega \in \mathbb{R}$. This corresponds, in both cases, to the *analysis* phase of the given signal f; the inversion formula (15) and the series expansion (20) then correspond to the *synthesis* or *reconstruction* phase of the signal.

Since $\gamma_k(t) = e^{2\pi i k t/\Omega}$ form an orthonormal basis of $L^2(\mathbb{T}_\Omega)$, Pareseval's identity gives us

$$\sum_{k\in\mathbb{Z}} |\langle f, \gamma_k\rangle|^2 = \sum_{k\in\mathbb{Z}} |\hat{f}[k]|^2 = \|f\|_2^2. \tag{22}$$

It is also important to state the following result (which should be compared with the result 3. in Theorem 2).

Lemma 1 (Riemann-Lebesgue) *If $f \in L^2(\mathbb{T}_\Omega)$, then its Fourier coefficients $\hat{f}[k]$ satisfy*

$$\lim_{|k|\to\infty} \hat{f}[k] = 0. \tag{23}$$

Note: The above result is also valid for functions in the wider class $L^1(\mathbb{T}_\Omega.)$

The analogies between the Fourier transforms and series can also be extended to results on convolutions, provided an appropriate definition for convolution is given. Given two functions $f, g \in L^1(\mathbb{T}_\Omega)$ we define its convolution as

$$f * g(t) = \int_{\mathbb{T}_\Omega} f(u)g(t-u)du.$$

We then have the following result (cf. Theorem 3).

Theorem 7 *Let $f, g \in L^1(\mathbb{T}_\Omega)$, with corresponding sequences $(\hat{f}[k])_{k\in\mathbb{Z}}$ and $(\hat{g}[k])_{k\in\mathbb{Z}}$ of Fourier coefficients. Then, $f * g \in L^1(\mathbb{T}_\Omega)$ and the sequence of its Fourier coefficients is the product of the two sequences $(\hat{f}[k])_{k\in\mathbb{Z}}$ and $(\hat{g}[k])_{k\in\mathbb{Z}}$, i.e.*

$$\widehat{f * g}\,[k] = \hat{f}[k]\,\hat{g}[k], \quad k \in \mathbb{Z}.$$

IV. Discrete Time Fourier Transform (DTFT)

The equality (22) shows that, given a function in $L^2(\mathbb{T}_\Omega)$, the sequence of its Fourier coefficients is in the space $\ell^2(\mathbb{Z})$. One can also "move" the other way around. Let $f = (f[k])_{k\in\mathbb{Z}}$ be a given sequence in $\ell^2(\mathbb{Z})$. Then, for any chosen $\Omega > 0$, the trigonometric series

$$\sum_{k\in\mathbb{Z}} f[k] e^{-2\pi i k\omega/\Omega} \tag{24}$$

converges (with respect to the $\|\cdot\|_2$ norm defined by (4)), to a certain function in the space $L^2(\mathbb{T}_\Omega)$. We call this function the *discrete time Fourier transform* (corresponding to Ω) of the sequence $f = (f[k])$ and denote it by $\hat{f}(\omega)$. That is, we have

$$\hat{f}(\omega) = \sum_{k\in\mathbb{Z}} f[k] e^{-2\pi i k\omega/\Omega}. \tag{25}$$

One can show that the Fourier coefficients of this function $\hat{f}$ are precisely the given numbers $f[k]$, that is, we have

$$\int_{\mathbb{T}_\Omega} \hat{f}(\omega) e^{2\pi i k\omega/\Omega} d\omega = f[k], \tag{26}$$

which can be seen as an inversion result. The equality (25) is also known to hold for almost all ω. Moreover, if the given sequence is known to decrease "faster" than just being in $\ell^2(\mathbb{Z})$, namely if $f = (f[k])_{k\in\mathbb{Z}} \in \ell^1(\mathbb{Z})$, then the series on the r.h.s. of (25) converges uniformly and defines a continuous function $\hat{f}(\omega)$, for all $\omega \in \mathbb{R}$.

If $f, g \in \ell^1(\mathbb{Z})$, we define the convolution $f * g$ of these two sequences by

$$(f * g)[k] := \sum_{l\in\mathbb{Z}} f[l] g[k-l]. \tag{27}$$

We again have a result concerning the behaviour of the (discrete) Fourier transform with respect to convolution.

Theorem 8 *Let $f, g \in \ell^1(\mathbb{Z})$ and let $\hat{f}, \hat{g}$ denote their discrete Fourier transforms. Then, $f * g \in \ell^1(\mathbb{Z})$ and*

$$\widehat{f * g}(\omega) = \hat{f}(\omega)\, \hat{g}(\omega) \tag{28}$$

V. Discrete Fourier Transform (DFT)

We now concentrate on the case where our signal is simultaneously discrete in time and finite, $f = (f[k])_{k=0}^{N-1}$. As already mentioned, we can also think of f as a periodic sequence $f = (f[k])_{k\in\mathbb{Z}}$ of period N (i.e. as an element in $\ell(\mathbb{Z}_N)$) by letting $f[k] = f[k \bmod N]$, for all $k \in \mathbb{Z}$.

It is easy to show that the set of N vectors $\gamma_k; k = 0, \ldots, N-1$, defined by

$$\gamma_k[n] := e^{2\pi ikn/N}; \; n = 0, \ldots, N-1, \tag{29}$$

is an orthogonal basis of $\ell(\mathbb{Z}_N)$ and that $\|\gamma_k\|^2 = N$. Hence, any signal $f \in \ell(\mathbb{Z}_N)$ admits the following expansion

$$f[n] = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}[k] e^{2\pi ikn/N}; \quad n = 0, 1, \ldots, N-1, \tag{30}$$

where the coefficients $\hat{f}[k]$ are given by

$$\begin{aligned} \hat{f}[k] &= \langle f, \gamma_k \rangle \\ &= \sum_{n=0}^{N-1} f[n] e^{-2\pi ikn/N}; \; k = 0, 1, \ldots, N-1. \end{aligned} \tag{31}$$

Formula (31) defines the so-called *discrete time Fourier series* or *discrete Fourier transform* (DFT) of f and formula (30) the *inverse discrete transform*. Naturally, the following *Parseval's identity* holds

$$\sum_{k=0}^{N-1} |f[k]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |\hat{f}[k]|^2.$$

Because of the N-periodicity of the functions $e^{-2\pi ikn/N}$, we can also see (31) as a function defined on $\mathbb{Z}_N$. This means that the discrete Fourier transform can be seen either as a map from $\mathbb{C}^N$ into $\mathbb{C}^N$ or as a map from $\ell(\mathbb{Z}_N)$ into $\ell(\mathbb{Z}_N)$. Let's introduce the following standard notation

$$W_N := e^{-2\pi i/N}. \tag{32}$$

Then, the discrete Fourier transform of $f = (f[n])_{n=0}^{N-1}$ can be defined by

$$\hat{f}[k] = \sum_{n=0}^{N-1} f[n] W_N^{kn}. \tag{33}$$

The discrete Fourier transform (as a linear transformation from $\mathbb{C}^N$ into $\mathbb{C}^N$) can also be defined using the $N \times N$ matrix (called the $N^{\underline{\text{th}}}$ *order DFT matrix*),

$$M = (m_{kn}), \quad m_{kn} = W_N^{kn}; \quad k, n = 0, \ldots, N-1.$$

It is simply given by

$$\hat{f} = Mf.$$

Given two sequences $f, g \in \ell(\mathbb{Z}_N)$, we define its convolution by

$$(f * g)[k] = \sum_{l=0}^{N-1} f[l] g[k-l], k = 0, 1, \ldots, N-1. \tag{34}$$

(Recall that the sequences are periodic of period N, i.e. $g[k] = g[k \bmod N]$.)

Once more, we have the usual property relating the Fourier transform of convolutions and the product of Fourier transforms.

Theorem 9 *Let $f, g \in \ell(\mathbb{Z}_N)$ and let $\widehat{f}$ and $\widehat{g}$ denote their DFT's. Then, we have*

$$\widehat{f * g}[k] = \hat{f}[k]\hat{g}[k].$$

• *Relation of DFT to Fourier coefficients*

Assume that we know the period Ω of a certain function f as well as N of its values $y[n] := f(t_n)$ at the equally spaced points

$$t_n := n\frac{\Omega}{N}; n = 0, 1 \ldots, N-1, \tag{35}$$

and that we want to make use this information to approximate the Fourier coefficients $\hat{f}[k]$ of f. In other words, we want to compute

$$\hat{f}[k] = \frac{1}{\Omega}\int_0^{\Omega} f(t)e^{-2\pi i k t/\Omega}dt. \tag{36}$$

If we approximate the integral in (36) by a left-endpoint, uniform Riemann sum, based on the points t_n, we obtain

$$\begin{aligned} \hat{f}[k] &\approx \frac{1}{\Omega}\sum_{n=0}^{N-1} f(t_n)e^{-2\pi i k t_n/\Omega} \times \frac{\Omega}{N} \\ &= \frac{1}{N}\sum_{n=0}^{N-1} y[n]e^{-2\pi i k n/N}. \end{aligned} \tag{37}$$

The above formula shows that the $k^{\underline{\text{th}}}$ Fourier coefficient of the function f is approximately given by $\frac{1}{N}\hat{y}[k]$, where $(\hat{y}[k])_{k=0}^{N-1}$ is the N-point discrete Fourier transform of the vector $(y[n])_{n=0}^{N-1} = (f(n\frac{\Omega}{N}))_{n=0}^{N-1}$.

Note: The approximation described by the formula (37) has to be interpreted very carefully. Note that the r.h.s of (37) has period N in the variable k and the same is not true for the sequence $(\hat{f}[k])$ (typically, $\hat{f}[k] \to 0$, as $k \to \infty$). The approximation (37) will usually be used only to calculate coefficients $\hat{f}[k]$ for $|k| << N$, e.g. for $|k| \leq N/8$; for a justification of this "rule of thumb", see e.g. [47].

VI. Transforms in several dimensions

All the transforms referred so far were given for the one-dimensional case, i.e. for functions of a single variable. The extension of these transforms to higher dimensions is straightforward. For example, the Fourier transform of a function $f \in L^1(\mathbb{R}^d)$ is defined by

$$\hat{f}(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} f(\boldsymbol{x})e^{-2\pi i\langle\boldsymbol{\omega},\boldsymbol{x}\rangle}d\boldsymbol{x}, \quad \boldsymbol{\omega} \in \mathbb{R}^d. \tag{38}$$

In the particular case of dimension $d = 2$ (this is of special importance due to its applications in image processing), we have

$$\hat{f}(\omega, \xi) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) e^{-2\pi i(\omega x + \xi y)} dx\, dy, \quad (\omega, \xi) \in \mathbb{R}^2. \tag{39}$$

The evaluation of the Fourier transform of a 2D-function is especially simple when the function is separable, i.e. can be written as

$$f(x, y) = g(x)h(y).$$

In that case, its Fourier transform is simply given by

$$\hat{f}(\omega, \xi) = \hat{g}(\omega)\hat{h}(\xi),$$

where $\hat{g}$ and $\hat{h}$ are the one-dimensional transforms of g and h. The basic transformational properties of a d-dimensional Fourier transform are essentially the same as in one dimension, with one new feature: the Fourier transform commutes with rotations, i.e. if R denotes a rotation in $\mathbb{R}^d$, then

$$f(R\boldsymbol{x}) \longrightarrow \hat{f}(R\boldsymbol{\omega}).$$

- *Fourier transforms of radial functions*

The fact that the Fourier transform commutes with rotations has the following interesting consequence. A function f defined in $\mathbb{R}^d$ is called *radial* if $f(R\boldsymbol{x}) = f(\boldsymbol{x})$ for all rotations R, i.e. $f(\boldsymbol{x})$ depends only on $|\boldsymbol{x}|$, where we used the simplified notation $|\cdot|$ for the Euclidean norm$\|\cdot\|_2$ in $\mathbb{R}^d$. If f is radial – say $f(\boldsymbol{x}) = g(|\boldsymbol{x}|)$ – then so is its Fourier transform – $\hat{f}(\boldsymbol{\omega}) = h(|\boldsymbol{\omega}|)$, say. In this case the integral formula relating f and $\hat{f}$ can be written in polar coordinates to yield h directly in terms of g. Let us illustrate with the two-dimensional case. With $\boldsymbol{x} = r(\cos\theta, \sin\theta)$ and $\boldsymbol{\omega} = \rho(\cos\phi, \sin\phi)$, we have $\langle x, \omega\rangle = r\rho\cos(\theta - \phi)$, and hence

$$\begin{aligned}
\hat{f}(\omega) &= \int_{\mathbb{R}^2} f(\boldsymbol{x}) e^{-2\pi i\langle x, \omega\rangle} d\boldsymbol{x} \\
&= \int_0^\infty \int_0^{2\pi} g(r) e^{-2\pi i r\rho\cos(\theta-\phi)} r d\theta dr \\
&= \int_0^\infty g(r) \left[\int_0^{2\pi} e^{-2\pi i\rho r\cos\theta} d\theta\right] r dr
\end{aligned}$$

By recalling the definition of the zero-order Bessel function of the first kind

$$J_0(z) = \frac{1}{2\pi} \int_0^{2\pi} e^{-iz\cos\theta} d\theta,$$

we obtain

$$h(\rho) = 2\pi \int_0^\infty g(r) J_0(2\pi\rho r) r dr. \tag{40}$$

The integral on the r.h.s (without the factor 2π) is called the *Hankel transform of order zero* of g.

- *Projection*

Suppose that we project a two-dimensional function $f(x, y)$ onto the x-axis, i.e we form

$$p(x) = \int_{\mathbb{R}} f(x, y)dy$$

Then, the (one-dimensional) Fourier transform of p is

$$\begin{aligned} \hat{p}(\omega) &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(x, y)dye^{-2\pi i\omega x} \right] dx \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y)e^{-2\pi i(\omega x+0y)}dxdy = \widehat{f}(\omega, 0) \end{aligned}$$

So, the transform of the projection of $f(x, y)$ onto the x-axis is $\widehat{f}(\omega, \xi)$ evaluated along the ω-axis. This, together with the rotation property, implies that the Fourier transform of the projection onto a a line at an angle θ with the x-axis is just the Fourier transform computed along a line at an angle θ with the ω-axis. This projection property can be used e.g. in computerized axial tomography; see, e.g. [13].

VII. Fourier Transform on Groups

It is possible to give a unified view of all of the different Fourier transforms described above. This is done by considering them as particular cases of a more general theory of *Fourier transforms on groups*. To present this theory in full detail requires ideas from topology and measure theory which are beyond the scope of these notes. We will, however, try to give a very brief idea of the main points (for simplicity, we will concentrate in the 1-D case).

A. Groups, Subgoups, Cosets

We start by recalling the notion of a *group*. A set G forms a *group* with respect to a certain binary operation $\oplus$, if the following properties hold:

1. *Closure* $\forall f, g \in G, \quad f \oplus g \in G$
2. *Associativity* $\forall f, g, h \in G, \quad (f \oplus g) \oplus h = f \oplus (g \oplus h)$
3. *Identity* $\exists 0_G \in G : \forall g \in G \quad 0_G + g = g + 0_G = g$
4. *Inverse* $\forall g \in G \exists -g \in G : g \oplus -g = -g \oplus g = 0_G$

As examples of groups especially important for our work, we have:

1. the set of real numbers $\mathbb{R}$, under addition;
2. the set of integers $\mathbb{Z}$, under addition;
3. the set $N\mathbb{Z}$, N fixed integer, under addition;
4. the set $\Omega\mathbb{Z}$, $\Omega > 0$, under addition;
5. the unit circle S^1 in the complex plane (i.e. the set of complex numbers of modulus 1), under multiplication.

A group is called Abelian if the operation $\oplus$ is commutative, i.e. $f \oplus g = g \oplus f$, for all $f, g \in G$.

A group G is *locally compact* if it has a topological structure such that the map $(f, g) \to f \oplus -g$ is continuous and every point in G has a compact neighbourhood. The group $\mathbb{R}$ is naturally a locally compact group (with the usual topology on $\mathbb{R}$). In fact, all the groups referred to in our examples are locally compact Abelian (LCA) groups.

A *subgroup* K of G is a subset of G which is also a group with respect to the same group operation. We use the notation $K \leq G$ (respectively $K < G$) to indicate that H is a subgroup of G (not equal to G itself). For example, for any N, $N\mathbb{Z}$ is a subgroup of $\mathbb{Z}$; the integers $\mathbb{Z}$ also form a subgroup of the additive group $\mathbb{R}$.

If $K < G$ and $g \in G$, we define the *coset* $g \oplus K$ of K in G as the set

$$g \oplus K = \{g + k : k \in K\}.$$

If G is an Abelian group with subgroup $K < G$, then the set of all cosets of K in G is a group under the following operation inherited from $\oplus$ (for which we use the same symbol $\oplus$):

$$(f \oplus K) \oplus (g \oplus K) := (f \oplus g) \oplus K. \tag{41}$$

This group is denoted by G/K (the *quotient group* of G modulo K).

It is easy to see that the group $\mathbb{Z}/N\mathbb{Z}$ is finite and has exactly N distinct elements. A set of *coset representatives* of G/K is a set S of elements of G such that every coset in G/K contains exactly one element of S. For example, a set of coset representatives of $\mathbb{Z}/N\mathbb{Z}$ can be taken to be $\{0, 1, \ldots, N-1\}$. When we use coset representatives instead of writing the full coset notation itself, we must remember that the operation involved is *modular*. In this sense, we can identify the group $\mathbb{Z}/N\mathbb{Z}$ with the group formed by the set $\{0, 1, \ldots, N-1\}$ with the operation of addition *modulo* N. Similarly, the group $\mathbb{T}_\Omega := \mathbb{R}/\Omega\mathbb{Z}$ can be identified with the group whose set of elements is $[0, \Omega)$ (or any other interval of length Ω) and whose operation is addition modulo Ω.

Let G and H be two groups with operations $\oplus_G$ and $\oplus_H$, respectively. A *homomorphism* from G to H is a map $\phi : G \to H$ such that

$$\phi(f \oplus_G g) = \phi(f) \oplus_H \phi(g), \quad \forall f, g \in G.$$

Is the homomorphism is bijective, we call it an *isomorphism.* For example, the function $\phi : \mathbb{T}_\Omega \to S^1$ defined by

$$\phi(t + \Omega\mathbb{Z}) = e^{2\pi i t/\Omega}, \quad t \in \mathbb{R}$$

is an isomorphism between the additive group $\mathbb{T}_\Omega$ and the multiplicative group S^1. In some sense, we can view the two groups S^1 and T_Ω as the *same* group. Another important example of two isomorphic groups is given by the group $\mathbb{Z}/N\mathbb{Z}$ and the group S_N of all the $N^{\underline{\text{th}}}$ roots of unity,

$$S_N := \{1, W_N, W_N^2, \ldots, W_N^{N-1}\}, \quad W_N := e^{2\pi i/N},$$

under multiplication. An isomorphism between the two groups is given by

$$\phi(k + N\mathbb{Z}) = e^{2\pi i k/N} = W_N^k.$$

B. Characters of a group

For any Abelian group G, a *character* γ of G is a homomorphism of G into the group S^1. The set of all *continuous* characters of G is denoted by $\hat{G}$ and is itself an Abelian group under the operation of pointwise multiplication. This is called the *dual* group of G.

For example, when G is the additive group $\mathbb{R}$, one can show that the characters are the functions

$$\gamma_\omega(x) = e^{2\pi i \omega x}, \quad x \in \mathbb{R},$$

for all $\omega \in \mathbb{R}$.

Note: The choice of the exponent 2π associated with the "index" ω is just for convenience; any other real number would do, which would correspond to a simple renaming of the functions γ.

This is easily seen to be isomorphic to $\mathbb{R}$ itself (the mapping $\omega \mapsto \gamma_\omega$ defining an isomorphism). In that sense, we say tat $\mathbb{R}$ is self-dual and write $\hat{\mathbb{R}} = \mathbb{R}$. One can also identify the characters of the group $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$: they are the functions $\gamma_k; k = 0, 1 \ldots, N-1$ defined by

$$\gamma_k[m] = e^{2\pi i k m/N}, \quad m \in \{0, \ldots, N-1\}.$$

The function $\phi : k \mapsto \gamma_k$ is an isomorphism between $\mathbb{Z}_N$ and $\hat{\mathbb{Z}}_N$ and, in that sense, $\mathbb{Z}_N$ is also self-dual.

Finally, we describe the characters of the group $\mathbb{T}_\Omega$. They are the functions

$$\gamma_k(t) := e^{2\pi i k t/\Omega}, \quad t \in [0, \Omega),$$

for all $k \in \mathbb{Z}$. The dual group of $\mathbb{T}_\Omega$ is thus isomorphic to $\mathbb{Z}$, the mapping $\phi : k \mapsto \gamma_k$ defining an isomorphism.

Since the dual group is also an Abelian group it is possible to define its set of characters, i.e. to define its dual $\hat{\hat{G}}$. It turns out that this group is always isomorphic to G. Hence, the dual group of $\mathbb{Z}$ is (isomorphic) to the Ω-torus $\mathbb{T}_\Omega$ (the particular choice of $\Omega > 0$ is not important, since all these groups are isomorphic to each other).

C. Integration on groups and Fourier transform

In order to be able to give an adequate definition of the Fourier transform on a (LCA) group, one has to introduce an appropriate concept of measure and corresponding integration on the group. It can be shown that in every LCA group, there exists a non-negative and regular measure, which is not identically zero and is *translation invariant*. This measure is unique (up to the multiplication by a positive constant) and is called the *Haar measure* of G. Integration on the group G will always be understood with respect to such measure. For the construction of such a measure, see e.g. [29] or [35]. In our cases, we simply refer that this measure is:

1. the usual Lebesgue measure, for the cases $G = \mathbb{R}$ and $G = \mathbb{T}_\Omega$ (with the "normalization" constant $\frac{1}{\Omega}$ in the latter case) ;
2. the usual counting measure for the discrete cases $G = \mathbb{Z}$ and $G = \mathbb{Z}/N\mathbb{Z}$.

Having defined integration on G, we can also introduce, in a natural way, the $L^p(G)$ spaces.

We can then define the *Fourier transform* of any function $f \in L^1(G)$: it is the function $\widehat{f}$, defined on $\widehat{G}$ by

$$\widehat{f}(\gamma) = \int_G f(t)\overline{\gamma(t)}dt, \quad \gamma \in \widehat{G}. \tag{42}$$

Note that the Fourier transform of a function defined on G is actually a function defined on $\hat{G}$. This means, for example, that in the case of $G = \mathbb{R}$, we should have written the Fourier transform of a function f as $\hat{f}(\gamma_\omega)$. However, due to the identification of $\mathbb{R}$ with $\hat{\mathbb{R}}$, this is naturally shortened to $\hat{f}(\omega)$.

Having identified previously the characters $\gamma \in \widehat{G}$ for all the cases $G = \mathbb{R}$, $G = \mathbb{T}_\Omega$, $G = \mathbb{Z}$ and $G = \mathbb{Z}_N$, it is now simple to verify that the definitions (13), (21) , (25) and (31) all fit into this framework.

We also have an inversion theorem (cf. formulae (15), (20), (26) and (30)).

Theorem 10 *Let $f \in L^1(G)$ be such that $\hat{f} \in L^1(\hat{G})$. If the Haar measure of G is fixed, the Haar measure of $\hat{G}$ can be normalized so that the following inversion formula holds*

$$f(t) = \int_{\hat{G}} \hat{f}(\gamma)\gamma(t)d\gamma, \quad t \in G.$$

If we define the convolution of any two functions $f, g \in L^1(G)$ by

$$f * g(t) = \int_G f(u)g(t-u)du$$

we have the following result from which the results of theorems 3,7,8 and 9 are specific examples:

Theorem 11 *Let $f, g \in L^1(G)$ and let $\widehat{f}, \widehat{g}$ be their Fourier transforms. Then, $f * g \in L^1(G)$ and*

$$\widehat{f * g}(\gamma) = \hat{f}(\gamma)\hat{g}(\gamma).$$

Finally, we would like to remark that there is a natural way of extending the definition of the Fourier transform from $L^1(G)$ to $L^2(G)$ and that the Parseval formula holds

$$\int_G f(t)\overline{g(t)}dt = \int_{\widehat{G}} \hat{f}(\gamma)\hat{g}(\gamma)d\gamma.$$

For more details on this fascinating topic of Fourier transforms on groups see, e.g. [40]. We now turn to the problem of describing efficient algorithms for the computation of Fourier transforms.

VIII. Fast Fourier Transform

The Fast Fourier Transform – the most valuable algorithm of our lifetime.

Strang, 1993

A direct calculation of a N-point DFT requires $(N-1)^2$ multiplications and $N(N-1)$ additions, i.e. it involves a number of operations of order $O(N^2)$. For large N, this can be extremely time consuming. In 1965, Cooley and Tukey [19] proposed an algorithm to compute the DFT reducing the number of operations involved to $O(N \log_2 N)$, when $N = 2^r$. This algorithm, which has become known as the *Fast Fourier Transform*, had a tremendous impact and is responsible for the widespread use of DFT's in almost all branches of scientific computation, with particular emphasis on digital signal processing.

Note: In fact, as referred in [32], the basic idea of the FFT had already been discovered by Gauss, in 1805, as an efficient means of interpolating asteroid orbits. However, it was the Cooley and Tukey publication which popularized the use of the discrete Fourier transform; see [18].

Many variants of the basic FFT algorithm have also appeared subsequently. Here, we will briefly describe one of the most widely used of these algorithms, the so-called *decimation in time, radix 2 FFT*; for other variants the reader is referred to, e.g. [47], [11] or [24]. FTT programs in various computer languages can be found in [39]. The article by Burrus [12] gives an excellent summary and contains an extensive list of references on efficient algorithms to compute the DFT. A compiled bibliography on this topic (with more than 3400 entries!) is given in [41].

A. Decimation in Time Radix 2 FFT

We will assume that N is a power of 2, say $N = 2^r$, where r is a positive integer. Let us start by recalling the formula for the N-point DFT transform of a sequence $f = (f[k])_{k=0}^{N-1}$,

$$\hat{f}[n] = \sum_{k=0}^{N-1} f[k] W_N^{kn}, \tag{43}$$

where $W_N = e^{2\pi i/N}$. For simplicity, we will introduce the notation $F_n := \hat{f}[n]$. We can halve the N-point DFT in (43) in two sums, each of which is a $N/2$-point DFT:

$$\begin{aligned} F_n &= \sum_{k=0}^{N/2-1} f[2k] W_N^{2kn} + \sum_{k=0}^{N/2-1} f[2k+1] W_N^{2kn} W_N^n \\ &= \sum_{k=0}^{N/2-1} f[2k] W_{N/2}^{kn} + \sum_{k=0}^{N/2-1} f[2k+1] W_{N/2}^{kn} W_N^n \end{aligned}$$

We can thus write

$$\left.\begin{aligned} F_n &= F_n^0 + W_N^n F_n^1 \\ F_{n+N/2} &= F_n^0 - W_N^n F_n^1 \end{aligned}\right\} ; n = 0, \ldots, N/2-1, \tag{44}$$

where, for $j = 0, 1$,

$$F_n^j = \sum_{k=0}^{N/2-1} f[2k+j] (W_{N/2})^{kn}, \tag{45}$$

and where we have used the fact that $W_N^{j(N/2)+n} = \pm W_N^n$, depending on wether $j = 0, 1$. The DFT $(F_n)_{n=0}^{N-1}$ written in terms of the calculations (44)-(45) can be visualized as

$$\begin{array}{c} F_n^0 \longrightarrow F_n^0 + W_N^n F_n^1 \\ \times \\ F_n^1 \longrightarrow F_n^0 - W_N^n F_n^1 \end{array} \tag{46}$$

where $n = 0, 1, \ldots, N/2-1$. This diagram is called a *butterfly*. The butterfly (46) can be viewed as a construction of the DFT in $\mathbb{Z}_N$ in terms of two DFTs, F^0 and F^1, on $\mathbb{Z}_{N/2}$. In the same way, each F^0 and F^1 can be constructed in terms of a pair of two of DFTs on $\mathbb{Z}_{N/4}$. For example,

$$F_n^0 = F_n^{00} + W_{N/2}^n F_n^{01}$$

and

$$F_{n+N/4}^0 = F_n^{00} - W_{N/2}^n F_n^{01},$$

for $n = 0, 1, \ldots, N/4-1$, where F^{00} is the DFT of $(f[0], f[4], \ldots, f[N-4])$ and F^{01} is the DFT of $(f[2], f[6], \ldots, f[N-2])$. Since $N = 2^r$, this procedure can be repeated and after $r = \log_2 N - 1$ stages we reach a point where we are performing $N/2$ 2-point DFTs, which consist of adding and subtracting two points. Computationally, it is convenient to compute the 2-point DFTs first, then the 4-point DFts, etc.

B. Bit Reversal

Let $f : \mathbb{Z}_N \to \mathbb{C}$ be given and suppose we want to compute the DFT F in the natural ordering $(F_0, \ldots, F_N)$. From (45), it is clear that if we begin with the DFTs of the pairs $(f[0], f[1]), (f[2], f[3]), \ldots$ we will not obtain F in the natural ordering. For example, when $N = 8$, the input indices must be ordered as $(0, 4, 2, 6, 1, 5, 3, 7)$ so that the output sequence will appear in the natural order. This ordering is obtained by *bit reversal*. Bit reversal (at *level* r or of *order* r) is defined recursively as follows. For $r = 1$, the bit reversal ordering (of the set $\{0, 1\}$) is the ordered pair $(0, 1)$. At level $r; r = 2, 3, \ldots$, the bit reversal ordering of the set $\{0, 1, \ldots, 2^r - 1\}$ is the 2^r-tuple

$$(2b_0, \ldots, 2b_{M-1}, 2b_0 + 1, \ldots, 2b_{M-1} + 1), \tag{47}$$

where $M = 2^{r-1}$ and $(b_0, b_1, \ldots, b_{M-1})$ is the bit reversal ordering at level $r-1$. For example, bit reversal orderings at levels 2 and 3 are $(0, 2, 1, 3)$ and $(0, 4, 2, 6, 1, 5, 3, 7)$, respectively. The term bit reversal comes from the following observation. If $k \in \{0, 1, \ldots, 2^r - 1\}$ has the binary expansion

$$k = \sum_{j=0}^{r-1} \epsilon_j 2^j$$

then the number in the position $k; k = 0, \ldots, 2^{r-1}$ of the bit reversal ordering is obtained by "reversing" the order of the coefficients ϵ_j in the above expansion. It is important to observe that there are efficient algorithms for obtaining bit-reversed indices; see e.g. [11] or [47]. This last reference also describes efficient ways of performing the butterfly calculations involved in each step of the FFT algorithm.

IX. Fourier Related Transforms

A. Cosine and Sine Transforms

A.1 Fourier Sine and Cosine Transform

Using Euler's formula, we can write the Fourier transform of f as

$$\begin{aligned}\hat{f}(\omega) &= \int_{\mathbb{R}} f(t) e^{-2\pi i \omega t} dt \\ &= \int_{\mathbb{R}} f(t) \cos(2\pi\omega t) dt - i \int_{\mathbb{R}} f(t) \sin(2\pi\omega t) dt \\ &:= \mathcal{C}f(\omega) - i\mathcal{S}f(\omega),\end{aligned} \tag{48}$$

where $\mathcal{C}f(\omega)$ and $\mathcal{S}f(\omega)$ are called, the *Fourier cosine transform* and *Fourier sine transform* of f, respectively. Observe that if the function f is real-valued, then its Fourier transform can found by evaluating two real integrals. Also, if f is an even function, then the Fourier transform of f is simply its Fourier cosine transform and can be computed simply

$$\begin{aligned}\hat{f}(\omega) &= \int_{\mathbb{R}} f(t)\cos(2\pi\omega t)dt \\ &= 2\int_0^\infty f(t)\cos(2\pi\omega t)dt.\end{aligned}$$

Similarly, if f is an odd function,

$$\begin{aligned}\hat{f}(\omega) &= -i\int_{\mathbb{R}} f(t)\sin(2\pi\omega t)dt \\ &= -2i\int_0^\infty f(t)\sin(2\pi\omega t)dt.\end{aligned}$$

A.2 Fourier Sine and Cosine Series

It is simple to establish that the set of functions

$$\gamma_n(t) := \cos(\pi n t/\Omega), \quad n \in \mathbb{N}_0,$$

is an orthogonal basis of the space $L^2(\mathbb{T}_\Omega)$, and that $\|\gamma_0\|_2^2 = 1$ and $\|\gamma_n\|_2^2 = 1/2,\ n \in \mathbb{N}$. Hence, every function $f \in L^2(\mathbb{T}_\Omega)$ admits an expansion

$$f(t) = \frac{1}{2}A_0 + \sum_{n=1}^{\infty} A_n \cos(\pi n t/\Omega)\,, \tag{49}$$

where

$$A_n = \frac{2}{\Omega}\int_0^\Omega f(t)\cos(\pi n t/\Omega)dt. \tag{50}$$

The series (49), with the coefficients given by (50), is called the *Fourier cosine series of f*.

Note: The above series is, in fact, the Fourier series of the *even* extension of f to $L^2(\mathbb{T}_{2\Omega})$.

In a similar manner, we define the Fourier sine series of f:

$$f(t) = \sum_{n=1}^{\infty} B_n \sin(\pi n t/\Omega)dt,$$

where

$$B_n = \frac{2}{\Omega}\int_0^\Omega f(t)\sin(\pi n t/\Omega)dt.$$

A.3 Discrete Cosine Transforms

There are also discrete versions of the sine and cosine transforms. Here, we refer to four established discrete cosine transforms (DCT-I through DCT-IV). The (two-dimensional version) of DCT-II and DCT-IV are constantly applied in image processing and have a FFT implementation, which makes them especially useful. The DCTs (in fact DCT-II) was only discovered in 1974, [1]. All four types of DCT are orthogonal transforms and use bases for the space $\mathbb{C}^N$ (or $\ell(\mathbb{Z}_N)$) that involve only *cosines.* For $k, n = 0, 1, \ldots, N-1$ the nth component of the kth basis vector is

DCT-I	$\cos\left(nk\frac{\pi}{N-1}\right)$	(divide by $\sqrt{2}$ when $k, n = 0, N-1$)
DCT-II	$\cos\left((n+\frac{1}{2})k\frac{\pi}{N}\right)$	(divide by $\sqrt{2}$ when $k = 0$)
DCT-III	$\cos\left(n(k+\frac{1}{2})\frac{\pi}{N}\right)$	(divide by $\sqrt{2}$ when $n = 0$)
DCT-IV	$\cos\left((n+\frac{1}{2})(k+\frac{1}{2})\frac{\pi}{N}\right)$	

If we consider the matrices $C_{\mathrm{I}}, C_{\mathrm{II}}, C_{\mathrm{III}}$ and C_{IV} whose columns are the above vectors, then each of the DCT-T transforms $\widehat{f}_{\mathrm{T}}$ of a vector $f \in \mathbb{C}^N$ is defined by

$$\hat{f}_{\mathrm{T}} = C_{\mathrm{T}}\, f; \quad \mathrm{T} = \mathrm{I}, \mathrm{II}, \mathrm{III}, \mathrm{IV}. \tag{51}$$

All vectors have norm $\sqrt{N/2}$; hence, we have, for example (using the DCT-IV transform), that any vector $f \in \mathbb{C}^N$ can be written as

$$f[n] = \frac{2}{N}\sum_{k=0}^{N-1} \hat{f}_{\mathrm{IV}}[k] \cos\left((n+\frac{1}{2})(k+\frac{1}{2})\frac{\pi}{N}\right)$$

where

$$\hat{f}_{\mathrm{IV}}[k] = \sum_{n=0}^{N-1} f[n] \cos\left((n+\frac{1}{2})(k+\frac{1}{2})\frac{\pi}{N}\right).$$

Similar expressions for the transform and corresponding inverse for the DCT-I – DCT-III are easily written.

B. Hartley Transform

The *Hartley transform* $\mathcal{H}f$ is obtained by combining the sine and cosine transforms replacing $-i$ by 1, i.e.

$$\begin{aligned}\mathcal{H}f(\omega) &= \mathcal{C}f(\omega) + \mathcal{S}f(\omega)\\ &= \int_{\mathbb{R}} f(t)\mathrm{cas}(2\pi\omega t)dt,\end{aligned} \tag{52}$$

where $\mathrm{cas}(t) := \cos t + \sin t$. The Hartley transform was initially proposed by Hartley in 1942 [31], but was virtually ignored until it was reintroduced by Bracewell [10] in 1983.

The Hartley transform has the advantage that is real-valued for a real-valued signal, but it lacks some of the important properties of the Fourier transform; a thorough investigation of Hartley transforms can be found in [9]. There is also a discrete version of the Hartley transform and fast algorithms for its computation (Fast Hartley Transform).

C. Laplace Transform

Fourier transforms were defined for *real* values of the frequency variable. A more general class of transforms can be obtained if the frequency variable is allowed to be complex.

We define the (bilateral) *Laplace transform* of a function f by

$$\mathcal{L}f(s) = \int_{\mathbb{R}} f(t)e^{-st}dt \tag{53}$$

where $s \in \mathbb{C}$. Note that, when $s = 2\pi i\omega$, $\mathcal{L}f(s) = \hat{f}(\omega)$ and so, as it might be expected, the Laplace transforms has many important properties similar to those of the Fourier transform. When $s = \sigma + 2\pi i\omega$, then $\mathcal{L}f(s)$ is the Fourier transform of $g(t) = f(t)e^{-i\sigma t}$, i.e. is the transform of an exponentially weighted signal.

Note: The more frequently used *unilateral* Laplace transform can be defined as the Laplace transform of $f(t)u(t)$, where $u(t)$ is the unit-step function defined by $u(t) = 1$, for $t \geq 0$ and $u(t) = 0$ otherwise.

The above transform does not, in general, converge for all values of s. The set of values for which (53) converges is called the region of convergence (ROC). The ROC has the following important poperties:

1. it consists of strips in the complex plane parallel to the to the $i\omega$ axis i.e. is of the form $A \leq \mathrm{Re}\,(s) \leq B$ where A and B may be $-\infty$ and $+\infty$, respectively; (In the extreme cases, the $\leq$ sign might have to be replaced by $<$);
2. if $f(t)$ is right-sided (left-sided), i.e. is zero for $t < T_0$ (i.e is zero for $t > T_1$), then $B = +\infty$ ($A = -\infty$).
3. if $f(t)$ is time-limited (i.e. $f(t) = 0$ for $T_0 < t < T_1$, then its ROC is the whole complex plane (provided it converges at some point);
4. if the $i\omega$ axis is contained in the ROC, then the Fourier transform of f exists.

The Laplace transform can be inverted. Its inverse is given by

$$f(t) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} e^{st}\mathcal{L}f(s)ds,$$

where σ is chosen inside the ROC.

The (unilateral) Laplace transform is particularly useful for solving initial value problems. For a comprehensive treatment of the Laplace transforms and its applications, we refer the reader to [22] or [5].

D. z-Transform

Just as the Laplace transform was a generalization of the Fourier transform, the *z-transform* can also be introduced as a generalization of the discrete time Fourier transform. For a given sequence $f = (f[k])_{k\in\mathbb{Z}}$, we define its z-transform as

$$Z(f[k]) := F(z) := \sum_{k\in\mathbb{Z}} f[k]z^{-k}, \tag{54}$$

where $z \in \mathbb{C}$. Again, the transform is only defined for the values of z for which the above series converges, these values defining its region of convergence (ROC). On the unit circle $z = e^{2\pi i\omega}$, this is the discrete-time Fourier transform ($\Omega = 1$), and for $z = \rho e^{2\pi i\omega}$, it is the discrete-time Fourier transform of the sequence $f[k]\rho^{-k}$. The ROC of the z-transform has properties "analogous" to the ROC of Laplace transforms:

1. it consists of a ring in the complex plane, i.e. is a set of the form $A \leq |z| \leq B$, where A may be zero and B may be $+\infty$. (In the extreme cases, the $\leq$ sign might have to be replaced by $<$).
2. if the sequence $f([k])$ is causal (i.e. $f[k] = 0$ for $k < 0$), then $B = +\infty$ ($\leq$ possibly replaced by $<$); if the sequence is anti-causal (i.e $f[k] = 0$ for $k > 0$), then $A = 0$ ($\leq$ possibly replaced by $<$);
3. if the sequence is of finite length and causal, the ROC is the entire plane, except possibly $z = 0$;
4. if the sequence is of finite length and anti-causal, the ROC is the entire z-plane except, possibly, the "point" $z = \infty$;
5. the discrete time Fourier transform of the sequence $f([k])$ converges absolutely if and only the ROC contains the unit circle.

The inverse z-transform involves the contour integration in the ROC and Cauchy's integral theorem. We have

$$f[n] = \frac{1}{2\pi i}\oint_C Z(f[k])z^{n-1}dz$$

where C denotes a contour around the origin lying in the ROC. The z-transform is very useful for the study of difference equations and discrete-time filters; more details can be seen, e.g. in [33] or [38].

E. Mellin Transform

The *Mellin transform* $\mathcal{M}f$ of a function f is defined by

$$\mathcal{M}f(z) = \int_0^\infty f(t)t^{z-1}dt. \tag{55}$$

If we make the change of variable $x = \log t$, we find that

$$\mathcal{M}f(z) = \int_\mathbb{R} f(e^x)e^{xz}dx \tag{56}$$

which shows that $\mathcal{M}f(-2\pi i\omega)$ is the Fourier transform (at ω) of the composition $f \circ \exp$; a good reference to read about Mellin transforms is the book by Bracewell [8].

F. Hilbert Transform

The *Hilbert transform* $\mathrm{H}f$ of $f \in L^2(\mathbb{R})$ is defined by

$$\mathrm{H}f(t) = \frac{1}{\pi}\int_{-\infty}^\infty \frac{f(u)}{t-u}du, \tag{57}$$

interpreting the integral as a Cauchy principal value, i.e. as

$$\lim_{\epsilon\to 0}\int_{|t-u|>\epsilon} \frac{f(u)}{t-u}\,du.$$

This transform is invertible, its inverse being simply $-H$, i.e.

$$f(t) = \frac{1}{\pi}\int_{-\infty}^\infty \frac{\mathrm{H}f(u)}{u-t}\,dt. \tag{58}$$

- *Analytic signals and Hilbert transform*

A function $f \in L^2(\mathbb{R})$ is said to be a (strong) *analytic signal* if its Fourier transform is zero for negative frequencies, i.e $\hat{f}(\omega) = 0$ for $\omega < 0$. If f is real valued, one can associate with f an analytic signal f_a in the following manner: f_a is the signal whose Fourier transform is given by

$$\hat{f}_a(\omega) = \begin{cases} 2\hat{f}(\omega), & \omega \geq 0 \\ 0, & \omega < 0. \end{cases} \tag{59}$$

One can show that if f_a if is the analytic signal associated with the real signal f, then $\mathrm{Re}\, f_a = f$ and $\mathrm{Im}\, f_a = \mathrm{H}f$, i.e.

$$f_a = f + i\mathrm{H}f. \tag{60}$$

Let

$$f_a(t) = A(t)e^{i\phi(t)}$$

The *envelope* $E(t)$ of the signal $f(t)$ is defined as $|A(t)| = \sqrt{f(t)^2 + \mathrm{H}f(t)^2}$ and $E(t)^2$ is the the so-called *instantaneous power*; the *instantaneous frequency* $\omega(t)$ is defined by $\omega(t) = \phi'(t)$. Hence, Hilbert transform analysis provides a method of determining the "instantaneous" frequency and power of a signal. This technique is widely used in communications systems analysis; see, e.g. [15].

G. *Haar and Walsh Transforms*

We consider again the space $L^2(\mathbb{T}_\Omega)$ and take, for simplicity, $\Omega = 1$, i.e. consider functions in the space $L^2[0,1]$. Besides the basis functions $\gamma_k(t) := e^{2\pi ikt}, \quad k \in \mathbb{Z}$, used in the Fourier series expansion, one may consider the use of other orthonormal bases for this space. We describe here two such bases, consisting of step functions.

Let $H(t)$ be the function defined by

$$H(t) = \begin{cases} 1, & 0 < t < \frac{1}{2} \\ -1, & 1/2 \leq t < 1 \end{cases} \tag{61}$$

This is called the *Haar function*. Then, the set of functions obtained by dyadic dilation and translation of this function, i.e.

$$H_{jk}(t) := 2^{j/2}H(2^j t - k), \ j \geq 0, \ k = 0, 1, \ldots, 2^j - 1, \tag{62}$$

together with the function $H_0 := \mathbf{1}_{[0,1)}$ (the characteristic function on the interval $[0,1)$), form an orthonormal basis for $L^2[0,1]$. Thus, every function $f \in L^2[0,1]$ admits a Haar series expansion

$$f(t) = f_H[0] + \sum_{j \geq 0} \sum_{k=0}^{2^j - 1} f_H[j,k] H_{jk}(t)$$

where the *Haar coefficients* $f_H[0]$ and $f_H[j,k]$ are given by

$$f_H[0] = \int_0^1 f(t)H_0(t)dt; \qquad f_H[j,k] = \int_0^1 f(t)H_{jk}(t)dt$$

To introduce the other basis consisting of step functions, we start by defining the so-called *Rademacher functions* r_n. For $n \geq 0$, consider the division of the interval $[0,1]$ into 2^n subintervals of equal length. Then, $r_n(t)$ is the function which takes the values $+1$ and -1, alternately, in each of these subintervals, starting with $+1$; in other words, if $d_n(t)$ is the $n^{\underline{\text{th}}}$ digit in the binary representation of t $(0 \leq t < 1)$, then

$$r_n(t) = (-1)^{d_n(t)}.$$

If $n \geq 0$ and $b_1, \dots, b_k$ are the digits in the binary representation of n, i.e. $n = (b_k \dots b_2 b_1)_2$, then the $n^{\underline{\text{th}}}$ *Walsh function* w_n is defined by

$$w_n(t) = r_1(t)^{b_1} r_2(t)^{b_2} \dots r_k(t)^{b_k}.$$

Then, the set of Walsh functions $\{w_n; n \geq 0\}$ is an orthonormal basis of $L^2[0,1]$. For an account of the applications of the Haar and Walsh functions in signal and image processing and other related fields see, e.g. [4]. Naturally, there are are also discrete versions of these transforms.

The Haar function is the first example (constructed by Haar in 1910 [28]) of an *orthogonal wavelet*, i.e. of a function $\psi \in L^2(\mathbb{R})$ whose dyadic dilations and translations $\psi_{jk} = 2^{j/2}\psi(2^j t - k)$; $j, k \in \mathbb{Z}$ constitute an orthonormal basis of $L^2(\mathbb{R})$; we will come back to this topic of wavelets in a little more detail in Section XI.

X. Windowed Fourier Transform

Recalling the expression for the Fourier transform

$$\hat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{-2\pi i \omega t} dt,$$

we see that $\hat{f}(\omega)$ depends on the values $f(t)$ for all time $t \in \mathbb{R}$. Hence, it is difficult to read any local behaviour of f from $\hat{f}$. In many applications, such as analysis of non-stationary signals or real time signal processing, the simple use of a Fourier transform may not be appropriate. In fact, one would like to dispose of an analytic tool that provides information both in time and frequency. One of the first ideas was simply to truncate the signal and to analyze only what happens on a finite interval $[-A, A]$. Mathematically, this corresponds to multiplying f by the characteristic function of this interval, $\mathbf{1}_{[-A,A]}$, and taking the Fourier transform of the product. We then have

$$\widehat{\mathbf{1}_{[-A,A]} f}(\omega) = (S_A * \hat{f})(\omega),$$

where $S_A(\omega) = \frac{\sin 2\pi A\omega}{\pi\omega}$. Thus, truncating the function results in convolving its spectrum with a cardinal sine. However, the cardinal sine decays slowly and has important lobes near the origin (hence there is poor localization in frequency). To avoid these problems, we can replace $\mathbf{1}_{[-A,A]}$ with more regular functions $W(t)$, called *windows*. Some typical choices include:

Bartlett or triangle window

$W(t) = (1 - \frac{|t|}{A})\mathbf{1}_{[-A,A]}$

Hamming and Hanning windows

$$W(t) = [\alpha + (1-\alpha)\cos(\pi t/A)]\mathbf{1}_{[-A,A]}$$

For $\alpha = 0.54$ we have Hamming's window and for $\alpha = 0.50$ we have Hanning's window.

Blackman window

$$W(t) = [0.42 + 0.5\cos(\pi t/A) + 0.08\cos(2\pi t/A)]\mathbf{1}_{[-A,A]}$$

Gaussian window

$$W(t) = Ce^{-\alpha t^2} \quad (C, \alpha > 0).$$

For more details and other choices of window functions, see e.g. [30] or [39]. All the windows described above are concentrated around the origin. We can then "slide" the window along the real axis and analyze the whole function. We then define the so-called *windowed Fourier transform* or *short time Fourier transform* (associated with the specific window W) as:

$$\mathcal{F}_W f(\omega, \tau) = \int_{\mathbb{R}} f(t)\overline{W}(t-\tau)e^{-2\pi i\omega t}dt. \tag{63}$$

Note: When a Gaussian window is used in the short time Fourier transform, this is usually referred as *Gabor transform*. If we define the family of functions $W_{\omega,\tau}$ by the result of two simple operations – translation by τ and modulation by ω – on the basic window W, i.e.

$$W_{\omega\tau}(t) := W(t-\tau)e^{2\pi i\omega t}, \tag{64}$$

we can view the windowed Fourier transform simply as the inner product of f with each of these functions:

$$\mathcal{F}_W f(\omega, \tau) = \langle f, W_{\omega,\tau}\rangle. \tag{65}$$

We then also have, by Plancherel formula,

$$\mathcal{F}_W f(\omega, \tau) = \langle \hat{f}, \hat{W}_{\omega,\tau}\rangle. \tag{66}$$

If W and $\hat{W}$ are localized around the origin, then $W_{\omega,\tau}$ is localized around the instant τ, while $\hat{W}_{\omega,\tau}$ is localized around the frequency ω. The value $\mathcal{F}_W f(\omega, \tau)$ thus provides an indication of how the function behaves around time τ and frequency ω.

The function f can always be recovered (in the L^2 sense), by a double integral

$$f(t) = \iint_{\mathbb{R}^2} \mathcal{F}_W f(\omega, \tau)W_{\omega\tau}(t)d\omega\, d\tau, \tag{67}$$

where we have assumed that the window W was chosen satisfying $\|W\|_2 = 1$. There is also an energy conservation property for the windowed Fourier transform:

$$\iint_{\mathbb{R}^2} |\mathcal{F}_W f(\omega, \tau)|^2 d\omega d\tau = \|f\|_2^2. \tag{68}$$

The windowed Fourier transform is even more familiar to signal analysis in its discrete version, where τ and ω are assigned regularly spaced values: $\tau = n\tau_0$ and $\omega = m\omega_0$, where $m, n \in \mathbb{Z}$ and $\tau_0, \omega_0 > 0$ are fixed. That is, we let

$$W_{m,n}(t) := e^{2\pi i m\omega_0 t} W(t - n\tau_0) \tag{69}$$

and compute the values

$$C_{m,n} = \langle f, W_{m,n} \rangle. \tag{70}$$

The question naturally arises of whether it is possible to reconstruct the given function f from its transform coefficients $C_{m,n}$ in a *numerically stable* way (i.e. in a manner not too "sensitive" to the unavoidable errors in the computed values). The answer is positive, provided the functions $W_{m,n}$ given by (69) constitute a *frame*, i.e. satisfy

$$A\|f\|_2^2 \leq \sum_{m,n\in\mathbb{Z}} |\langle f, W_{m,n}\rangle|^2 \leq B\|f\|_2^2, \tag{71}$$

for all $f \in L^2(\mathbb{R})$, with constants $0 < A \leq B < \infty$. The following theorem, whose proof can be seen, e.g. in [20], establishes necessary conditions on the parameters ω_0 and τ_0 for the set functions $\{W_{m,n} : m, n \in \mathbb{Z}\}$ to be a frame of $L^2(\mathbb{R})$.

Theorem 12 *Let $W \in L^2(\mathbb{R})$ be such that $\|W\|_2 = 1$. The windowed Fourier family $\{W_{m,n} : m, n \in \mathbb{Z}\}$ can only be a frame if*

$$\omega_0\tau_0 \leq 1. \tag{72}$$

The frame bounds A and B necessarily satisfy

$$A \leq \frac{1}{\omega_0\tau_0} \leq B. \tag{73}$$

In particular, a necessary condition for the functions (69) to be an orthonormal basis of $L^2(\mathbb{R})$ is that $\omega_0\tau_0 = 1$.

We also have the following important theorem, whose proof can again be seen in [20].

Theorem 13 (Balian-Low) *If $\|W\|_2 = 1$ and $\{W_{m,n} : m, n \in \mathbb{Z}\}$ is a windowed Fourier frame with $\omega_0\tau_0 = 1$, then*

$$\int_{\mathbb{R}} t^2 |W(t)|^2 dt = +\infty \quad \text{or} \quad \int_{\mathbb{R}} \omega^2 |\hat{W}(\omega)|^2 d\omega = +\infty.$$

This theorem shows, in particular, that we can not construct an ortogonal windowed Fourier basis with a differentiable window of compact support.

XI. Wavelet Transform

A. Continuous Wavelet Transform

The windowed Fourier transform computes the inner product of the function f with a family of functions $W_{\omega,\tau}$ obtained by translating and modulating the basic window W. The functions of this family are all of the "same size" (i.e. they all have the same spread in time and frequency). If the signal to be studied has components which are almost stationary associated with sudden variations, then the windowed Fourier analysis is not the appropriate tool, due the above fixed size of the windows. We now study a different transform which overcomes the above limitations, by using windows whose size naturally adjusts to frequencies. The idea of the *continuous wavelet transform* is again to compute the inner product of the function to be analyzed with a family of functions $\psi_{a,\tau}$ dependent on two parameters. In this case, however, these functions are obtained from a basic function (the *analyzing* or *mother wavelet*) by contractions or dilations (i.e. changes of scale)– controlled by the parameter a, and translations – controlled by the parameter τ. The mother wavelet ψ used for the analysis has to satisfy a certain technical condition, known as the *admissibility condition*. More precisely, we say that $\psi \in L^2(\mathbb{R})$ is a *wavelet* if it satisfies

$$C_\psi := \int_{\mathbb{R}} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < +\infty \tag{74}$$

In practice, we want to use a function ψ which behaves like a time window, i.e. we select ψ with a fast decay property in time (e.g. ψ and $t\psi(t) \in L^1(\mathbb{R})$). In this case, the admissibility condition (74) turns out to be equivalent to the condition

$$\int_{\mathbb{R}} \psi(t)dt = 0. \tag{75}$$

This indicates that ψ must "oscillate" above and below the t axis, i.e. must behave like a wave; this, together with the constraint that ψ decays fast (i.e. is "small") justifies the name *wavelet* adopted for these functions. Given a certain wavelet ψ (normalized so that $\|\psi\|_2 = 1$), we define the family of functions

$$\psi_{a,\tau}(t) := \frac{1}{\sqrt{|a|}} \psi(\frac{t-\tau}{a}); \quad a \in \mathbb{R}^* = \mathbb{R} \setminus \{0\}, \tau \in \mathbb{R}. \tag{76}$$

Then, the *continuous wavelet transform* (associated with the wavelet ψ) of f is defined by

$$\begin{aligned} \mathcal{W}_\psi f(a,\tau) &= \langle f, \psi_{a,\tau} \rangle \\ &= \frac{1}{\sqrt{|a|}} \int_{\mathbb{R}} f(t)\overline{\psi}(\frac{t-\tau}{a})dt, \ a \in \mathbb{R}^*, \tau \in \mathbb{R}. \end{aligned} \tag{77}$$

As in the windowed Fourier case, there is an inversion formula and a conservation of energy result, which can be stated as follows:

$$f(t) = \frac{1}{C_\psi} \iint_{\mathbb{R}^2} \mathcal{W}_\psi f(a,\tau) \psi_{a,\tau}(t) \frac{da\, d\tau}{a^2} \tag{78}$$

and

$$\frac{1}{C_\psi} \iint_{\mathbb{R}^2} |\mathcal{W}_\psi f(a,\tau)|^2 \frac{da\, d\tau}{a^2} = \|f\|_2^2. \tag{79}$$

If the function ψ is localized around $t = 0$ and $\hat{\psi}$ is localized around $\omega = 1$, then $\psi_{a,\tau}$ will be localized around τ whilst $\hat{\psi}_{a,\tau}$ will be localized around $\omega = \frac{1}{a}$. When $|a| > 1$ ($|\omega| = |\frac{1}{a}| < 1$), the function $\psi_{a,\tau}$ becomes a stretched version of ψ (less localized in time, more localized in frequency); on the contrary, when $|a| < 1$ ($|\omega| = |\frac{1}{a}| > 1$), $\psi_{a,\tau}$ will be a function more localized in time (a compressed version of ψ) and less localized in frequency; this is the already mentioned flexibility of the wavelet windows: their size naturally adjusts to the frequencies.

B. *Multiresolution Analysis (MRA)*

As usual, we might like to use a discretized version of the wavelet transform, i.e. to compute $W_\psi(a,\tau)$ only for a discrete set of values of a and τ. A very common choice is to take the dyadic points in the plane

$$a = 2^{-j}, \quad \tau = 2^{-j}k; \quad j,k \in \mathbb{Z}. \tag{80}$$

For the above choice of values, we thus consider the family of functions

$$\psi_{j,k}(t) := 2^{j/2}\psi(2^j t - k);\ j,k \in \mathbb{Z} \tag{81}$$

and compute the wavelet values

$$C_{j,k} = \langle f, \psi_{j,k} \rangle. \tag{82}$$

A natural challenge for the earlier researchers was to find ψ such that the corresponding set of functions (81) was an orthonormal basis of $L^2(\mathbb{R})$, in which case every function $f \in L^2(\mathbb{R})$ could be decomposed in a double series

$$f(t) = \sum_{j,k\in\mathbb{Z}} C_{j,k}\psi_{j,k}(t), \tag{83}$$

with the coefficients $C_{j,k}$ given by (82). A function with this property is called an *orthogonal wavelet*. In section VIII, we already mentioned the existence of one such function: the *Haar wavelet* (61). This is, however, a discontinuous function, and the converge of the series (83) is extremely slow. In the 80's, other orthogonal wavelets, with better properties, were discovered by J. O. Strömberg [42], Y. Meyer [37], G. Battle [3] and P. G. Lemarié [34].

These first constructions of wavelets seem a bit "miraculous"; Y. Meyer confesses "*I found my wavelets by trial and error; there was no underlying concept.*" In the end of 1986, Stéphane Mallat, in collaboration with Yves Meyer, introduce the important concept of *multiresolution analysis* (MRA). This structure gives a complete understanding of all the wavelet constructions obtained up to then, and allows the construction of new orthogonal wavelets. It is based on this concept, that I. Daubechies introduces a new class of wavelets (the so called *Daubechies wavelets*) which became of great importance in applications; these wavelets have important properties: they have compact support, are smooth (smoothness increasing with the size of support) and have a certain number of zero moments.

Another important consequence of the introduction of the AMR paradigm was the discovery of efficient computational algorithms for the decomposition and reconstruction of a function in a wavelet basis, the *fast wavelet transforms*.

A *multiresolution analysis* (MRA)(V_j, ϕ) of $L^2(\mathbb{R})$ is a sequence of closed subspaces of $L_2(\mathbb{R})$ and an associated function ϕ, called the *generator* or *scaling function*, satisfying:

1. $V_j \subset V_{j+1}, \quad \forall j \in \mathbb{Z}$
2. $\displaystyle\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$
3. $\displaystyle\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$
4. $v(t) \in V_j \iff v(2t) \in V_{j+1}$
5. The integer translates of ϕ, $\phi(t-k), k \in \mathbb{Z}$, form an orthonormal basis of the space V_0.

Note: The concept here introduced is sometimes referred as *orthogonal* AMR; in fact, Condition 5. can be replaced by the less stringent assumption that the $\phi(t-k)$ are a Riesz basis of V_0; in that case, an "orthogonalized" function $\phi^\perp$ such that $\phi^\perp(t-k)$ forms an o.n. basis of V_0 can always been obtained by a well-defined procedure; see, e.g. [20, pp. 139-140].

It follows from the properties of an AMR that, for each j, the set of functions $\{\phi_{j,k} := 2^{j/2}\phi(2^j. - k) : k \in \mathbb{Z}\}$ is an o.n. basis for the space V_j (the so-called nodal basis). Wavelets are associated with *detail spaces*, i.e. with complementary spaces W_j satisfying $V_{j+1} = V_j \oplus W_j$, where $\oplus$ denotes the orthogonal complement of V_j in V_{j+1}. The properties of the multiresolution analysis imply that $\bigoplus_{j\in\mathbb{Z}} W_j = L_2(\mathbb{R})$. Hence, if we can find a function ψ whose integer translates form an o.n. basis of W_0, then the collection $\{\psi_{j,k} := 2^{j/2}\psi(2^j \cdot -k) : j,\ k \in \mathbb{Z}\}$ will be an o.n. basis for the space $L_2(\mathbb{R})$ (a so-called *wavelet* *basis*), i.e. ψ will be an orthogonal wavelet. The basic principle of a multiresolution analysis is that ψ always exists and can be explicitly determined (from ϕ). In fact, we have the following theorem.

Theorem 14 *Let $(V_j)_{j\in\mathbb{Z}}$ be a MRA of $L^2(\mathbb{R})$ with scaling function ϕ. Then*

1. *there exists a sequence of scalars $(h_k) \in \ell^2(\mathbb{Z})$ such that*

$$\phi(t) = \sqrt{2}\sum_{k\in\mathbb{Z}} h_k \phi(2t-k) \tag{84}$$

2. *the function ψ defined by*

$$\psi(t) = \sqrt{2}\sum_{k\in\mathbb{Z}} g_k \phi(2t-k) \tag{85}$$

where the coefficients g_k are given by

$$g_k = (-1)^k \overline{h}_{1-k} \tag{86}$$

is an orthogonal wavelet., i.e. the set of functions $\{\psi_{j,k}(t) := 2^{j/2}\psi(2^j t - k),\ j,k \in \mathbb{Z}\}$ is an o.n. basis of $L^2(\mathbb{R})$.

Notes:

1. Equation (84), which is known as the *refinement equation* or the *two-scale equation* for the scaling function ϕ follows immediately by observing that $\sqrt{2}\phi(2t-k)$ is an o.n. basis of V_1 and hence the function $\phi \in V_0 \subset V_1$ must have a representation in that basis.

2. The sequence of coefficients $(h_k)_k \in \mathbb{Z}$ in (84) is called the *filter* of ϕ. These coefficients are, naturally, given by

$$h_k = \langle f, \phi_{1,k}\rangle = \sqrt{2}\int_{\mathbb{R}} f(t)\overline{\phi(2t-k)}dt. \tag{87}$$

3. tThere are other possible ways to define the coefficients g_k (in terms of h_k) so that (85) is an orthogonal wavelet; the different wavelets are, however, all closely related to each other; further details can be seen, e.g. in [20, pp. 135-136].

C. Fast Wavelet Transforms

We now show how the MRA structure leads to a very efficient iterative scheme for computing the coefficients of the expansion of a function f in a wavelet basis.

Let $(V_j)_{j\in\mathbb{Z}}$ be a MRA of $L^2(\mathbb{R})$, with scaling function ϕ and corresponding wavelet ψ. Properties 1. and 2. of the MRA show that any function $f \in L^2(\mathbb{R})$ can be arbitrarily well approximated by a function v_j in a certain space V_j, provided j is taken sufficiently large, i.e.

$$\forall\epsilon > 0 \quad \exists J \in \mathbb{Z} \quad \exists v_J \in V_J : \quad \|f - v_J\|_2 < \epsilon. \tag{88}$$

Let, as before, denote by W_j the orthogonal complement of V_j into V_{j+1} and let P_j and Q_j denote the orthogonal projectors of $L^2(\mathbb{R})$ into V_j and W_j, respectively; since $V_j \subset V_{j+1}$, we have that $Q_j = P_{j+1} - P_j$. Moreover, $P_jP_{j+1} = P_j$ and $Q_jP_{j+1} = Q_j$.

For each j, let v_j and w_j be the projections of f into V_j and W_j, respectively, i.e. let v_j and w_j be given by

$$v_j = P_jf \qquad \text{and} \qquad w_j = Q_jf. \tag{89}$$

We thus have,

$$\begin{aligned} v_J = P_Jf &= P_{J-1}f + (P_J - P_{J-1})f \\ &= v_{J-1} + w_{J-1} \\ &= v_{J-2} + w_{J-2} + w_{J-1} \\ &= \cdots = v_{J-M} + w_{J-M} + \cdots + w_{J-1}, \qquad M > 0, \end{aligned} \tag{90}$$

Property 2. of AMR ensures that, provided M is sufficiently large, one has

$$\|v_{J-M}\|_2 < \epsilon. \tag{91}$$

We can therefore conclude that any function in $L^2(\mathbb{R})$ can be reasonably well represented as a finite sum of functions belonging to the subspaces W_j and a remainder v_{J-M} in a space V_{J-M} which can be interpreted as a very coarse version of f. The decomposition (90) tells us the *details* that must be added to this *blurred* version of f to obtain the *fine* approximation v_J to f.

Let us assume that we know the approximation $v_J = P_Jf \in V_J$ to f and that we want to obtain the decomposition (90). Since, for every j, $\{\phi_{j,k} : k \in \mathbb{Z}\}$ and $\{\psi_{j,k} : k \in \mathbb{Z}\}$ are o.n. bases of V_j and W_j, respectively, to know the functions v_J and $v_{J-M}, w_{J-M}, \dots, w_{J-1}$, is equivalent to know their coefficients in these bases. Let $\boldsymbol{c}^j = (c_k^j)_{k\in\mathbb{Z}}$ be the sequence of the coefficients of $v_j = P_jf$ in the basis $\{\phi_{j,k} : k \in \mathbb{Z}\}$, i.e. let

$$c_k^j = \langle f, \phi_{j,k}\rangle, \quad k \in \mathbb{Z}, \tag{92}$$

and let $\boldsymbol{d}^j = (d_k^j)_{k\in\mathbb{Z}}$ be the sequence of the coefficients of $w_j = Q_jf$ in the basis $\{\psi_{jk} : k \in \mathbb{Z}\}$, i.e. let

$$d_k^j = \langle f, \psi_{jk}\rangle, \quad k \in \mathbb{Z}. \tag{93}$$

Hence, we aim to obtain the decomposition

$$v_J = \sum_{k\in\mathbb{Z}} c_k^{J-M}\phi_{J-M,k} + \sum_{j=J-M}^{J-1}\sum_{k\in\mathbb{Z}} d_k^j\psi_{j,k}. \tag{94}$$

Recall that ϕ satisfies the dilation equation, i.e. that

$$\phi(t) = \sum_{n\in\mathbb{Z}} h_n\phi_{1,n}(t)$$

Hence, we have

$$\begin{aligned}
\phi_{j-1,k}(t) &= 2^{(j-1)/2}\phi(2^{j-1}t-k)\\
&= 2^{(j-1)/2}\sum_{n\in\mathbb{Z}} h_n\phi_{1,n}(2^{j-1}t-k)\\
&= 2^{j/2}\sum_{n\in\mathbb{Z}} h_n\phi(2^jt-(2k+n))\\
&= \sum_{n\in\mathbb{Z}} h_n\phi_{j,2k+n}(t)\\
&= \sum_{n\in\mathbb{Z}} h_{n-2k}\phi_{j,n}(t).
\end{aligned} \tag{95}$$

Thus, one gets

$$\begin{aligned}
c_k^{j-1} &= \langle f,\phi_{j-1,k}\rangle\\
&= \langle f,\sum_{n\in\mathbb{Z}} h_{n-2k}\phi_{j,n}\rangle\\
&= \sum_{n\in\mathbb{Z}} \overline{h_{n-2k}}\,\langle f,\phi_{j,n}\rangle\\
&= \sum_{n\in\mathbb{Z}} \overline{h_{n-2k}}\,c_n^j.
\end{aligned} \tag{96}$$

In a totally similar manner, by making use of the equations (85) and (86), one gets

$$d_k^{j-1} = \sum_{n\in\mathbb{Z}} \overline{g_{n-2k}}\,c_n^j. \tag{97}$$

Starting from the sequence $\boldsymbol{c}^J = (c_n^J)$, formulae (96) and (97)above can be used, recursively, to obtain the sequences $\boldsymbol{c}^{J-M}, \boldsymbol{d}^{J-1}, \ldots, \boldsymbol{d}^{J-M}$, i.e. to obtain the desired decomposition for v_j; see the scheme in Figure 1.

The above transform can be easily inverted; starting from the sequences $\boldsymbol{c}^{J-M}, \boldsymbol{d}^{J-1}, \ldots, \boldsymbol{d}^{J-M}$, we can obtain the initial sequence of coefficients $\boldsymbol{c}^J$. We have, for each j,

$$\begin{aligned}
P_jf &= v_j = v_{j-1} + w_{j-1}\\
&= \sum_{l\in\mathbb{Z}} c_l^{j-1}\phi_{j-1,l} + \sum_{l\in\mathbb{Z}} d_l^{j-1}\psi_{j-1,l}
\end{aligned}$$

$$
\begin{tikzcd}
\boldsymbol{c}^{J} \arrow[r] \arrow[dr] & \boldsymbol{c}^{J-1} \arrow[r] \arrow[dr] & \boldsymbol{c}^{J-2} \arrow[r] & \cdots \arrow[r] \arrow[dr] & \boxed{\boldsymbol{c}^{J-M}} \\
 & \boxed{\boldsymbol{d}^{J-1}} & \boxed{\boldsymbol{d}^{J-2}} & & \boxed{\boldsymbol{d}^{J-M}}
\end{tikzcd}
$$

Figure 1: Decomposition Scheme

Therefore,

$$
\begin{aligned}
c_k^j &= \langle f, \phi_{j,k} \rangle \\
&= \langle P_j f, \phi_{j,k} \rangle \\
&= \sum_{l \in \mathbb{Z}} c_l^{j-1} \langle \phi_{j-1,l}, \phi_{j,k} \rangle + \sum_{l \in \mathbb{Z}} d_l^{j-1} \langle \psi_{j-1,l}, \phi_{j,k} \rangle. \qquad (98)
\end{aligned}
$$

But,

$$
\begin{aligned}
\langle \phi_{j-1,l}, \phi_{j,k} \rangle &= \langle \sum_{n \in \mathbb{Z}} h_{n-2l} \phi_{j,n}, \phi_{j,k} \rangle \\
&= \sum_{n \in \mathbb{Z}} h_{n-2l} \langle \phi_{j,n}, \phi_{j,k} \rangle \\
&= \sum_{n \in \mathbb{Z}} h_{n-2l} \delta_{n,k} = h_{k-2l}. \qquad (99)
\end{aligned}
$$

On the other hand,

$$
\langle \psi_{j-1,l}, \phi_{j,k} \rangle = \langle \sum_{n \in \mathbb{Z}} g_{n-2l} \phi_{j,n}, \phi_{j,k} \rangle = g_{k-2l}. \qquad (100)
$$

Hence, we get

$$
\begin{aligned}
c_k^j &= \sum_{l \in \mathbb{Z}} h_{k-2l} c_l^{j-1} + \sum_{l \in \mathbb{Z}} g_{k-2l} d_l^{j-1} \\
&= \sum_{l \in \mathbb{Z}} \left(h_{k-2l} c_l^{j-1} + g_{k-2l} d_l^{j-1} \right); \qquad (101)
\end{aligned}
$$

see the scheme in Fig.2.

$$
\begin{tikzcd}
\boldsymbol{d}^{J-M} \arrow[dr] & \boldsymbol{d}^{J-M+1} \arrow[dr] & & \boldsymbol{d}^{J-1} \arrow[dr] & \\
\boldsymbol{c}^{J-M} \arrow[r] & \boldsymbol{c}^{J-M+1} \arrow[r] & \cdots \arrow[r] & \boldsymbol{c}^{J-1} \arrow[r] & \boxed{\boldsymbol{c}^{J}}
\end{tikzcd}
$$

Figure 2: Reconstruction Scheme

Notes

1. Naturally, when implementing the algorithms, all the infinite sequences have to be truncated. Hence, when we apply the decomposition scheme, the initial sequence is always a finite sequence, i.e. a vector of certain length N, $(c_0^J, c_1^J, \ldots, c_{N-1}^J)$. Also, either the filter $(h_k)_{k\in\mathbb{Z}}$ is finite or, if we are working with wavelets that do not have compact support, will have to be truncated for a vector of a certain size L: $(h_{-m}, h_{-m+1}, \ldots, h_{-m+L-1})$.

2. Since the initial sequence has finite length it is necessary to know how to deal with the boundary points. For example, the formulae for c_0^{J-1} and $c_{N/2-1}^{J-1}$ are

$$c_0^{J-1} = \sum_{n=-m}^{-m+L-1} \overline{h_n} c_n^J$$

and

$$c_{N/2-1}^{J-1} = \sum_{n=-m+N-2}^{n=N+L-m-3} \overline{h_{n-N+2}} c_n^J$$

Hence, it is necessary to add m components at the left of the vector $\boldsymbol{c}^J$ and $L-2-m$ components at the end. This can be done in several ways; see, e.g. [36, pp. 282-290] for a discussion on different choices of these boundary conditions.

3. With an appropriate choice of the boundary conditions, formulae (96) and (97) show that in the first step of the decomposition we compute approximately $N/2$ coefficients c_k^{J-1} and $N/2$ coeficientes d_k^{J-1}. The next decomposition step is only applied to the coefficients c_k^{J-1} which represnt the part in V_{J-1} and so on. Hence, as the decomposition proceeds, less operations are involved If the filter length is L, the number of operations involved is of the order of

$$L \times \left(N + \frac{N}{2} + \frac{N}{4} + \cdots\right) < 2NL.$$

Hence, the number of operations involved in the fast wavelet transform is $O(N)$; cf. with $O(N \log N)$ for the FFT.

There are many important variants of the basic wavelet theory. Since it is impossible to present here a reasonable description (even very brief) of these variants, we just refer to some of these developments and indicate some references for the interested reader:

- Biorthogonal wavelets, introduced by Cohen, Daubechies and Feaveau in [14];

- *wavelet-packets*, introduced in [17] and applied in signal compression in [48]; we also recommend the book by Wickerhauser [49] and the article [46];
- Wilson bases– [21];
- local sine and co-sine bases – [16], [2];
- multiwavelets – [27];
- interpolatory wavelets – [23];
- lifting scheme and second generation wavelets – [43, 44, 45].

XII. Conclusion

The idea of transforming or decomposing an object (e.g. a function) in order to extract more "relevant" (for a specific purpose) information, and then reconstituting it, pervades all the areas of mathematics. This makes the subject of *mathematical transforms* extremely vast and impossible to cover, even in condensed form, in a set of notes. We were, therefore, forced to make a personal selection of topics. Our idea has been to focus on the most popular transforms, having also in mind their relevance in applied areas, such as signal processing.

We sincerely hope that these notes can be useful as a quick reference and a starting point for studying, more deeply, this fascinating area of mathematics.

References

[1] N. Ahmed, T. Natarajan, and K. R. Rao. Discrete cosine transform. *IEEE Trans. on Computers*, (C-23):90–93, 1974.

[2] P. Auscher, G. Weiss, and M. V. Wickerhauser. Local sine and cosine bases of Coifman and Meyer and the construction of smooth bases. In C. K. Chui, editor, *Wavelets: A Tutorial in Theory and Applications*, volume 2 of *Wavelet Analysis and Its Applications*, pages 181–216. Academic Press, Boston, 1992.

[3] G. Battle. A block spin construction of ondelettes. Part I: Lemarié functions. *Comm. Math. Phys.*, 110:601–615, 1987.

[4] K. G. Beauchamp. *Applications of Walsh and Related Functions.* Academic Press, London, 1984.

[5] R. E. Belmman and R. S. Roth. *The Laplace Transform.* World-Scientific, Singapore, 1984.

[6] J. J. Benedetto. *Harmonic Analysis and Applications.* CRC Press, New York, 1997.

[7] M. Bôcher. *Annals of Math.*, 7, 1906.

[8] R. Bracewell. *The Fourier Transform and its Applications*. McGraw-Hill, New York, 1965.

[9] R. Bracewell. *The Hartley Transform*. Oxford Press, New York, 1987.

[10] R. N. Bracewell. Discrete Hartley Transform. *J. Opt. Soc. Amer.*, 73:1832–1835, 1983.

[11] E. O. Brigham. *The Fast Fourier Transform and Its Applications*. Prentice-Hall, New Jersey, 1988.

[12] C. S. Burrus. Notes on the FFT. Department of Electrical and Computer Engineering, Rice University, September 1997.

[13] K. R. Castleman. *Digital Image Processing*. Prentice-Hall, New Jersey, 1996.

[14] A. Cohen, I. Daubechies, and J.-C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 45:486–560, 1992.

[15] L. Cohen. *Time-Frequency Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey, 1995.

[16] R. Coifman and Y. Meyer. Remarques sur l'analyse de Fourier à fenêtre. *C. R. Acad. Sci.*, 4312:259–261, 1991.

[17] R. Coifman, Y. Meyer, S. Quake, and M. V. Wickerhauser. Signal processing and compression with wavelt packets. In Y. Meyer and S. Roques, editors, *Proceedings of the International Conference on Wavelets, Marseille*, Paris, 1989. Masson.

[18] J. W. Cooley. How the FFT gained acceptance. *IEEE Signal Processing Magazine*, 9:10–13, 1990.

[19] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. of. Comput.*, 19(90):297–301, 1965.

[20] I. Daubechies. *Ten Lectures on Wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, 1992.

[21] I. Daubechies, S. Jaffard, and J.-L. Journé. A simple Wilson orthonormal basis with exponential decay. *SIAM J. of Math. Anal.*, 22:554–572, 1991.

[22] G. Doetsch. *Introduction to the Theory and Application of the Laplace Transform*. Springer-Verlag, Berlin, 1974.

[23] D. L. Donoho. Interpolating wavelet transforms. *Appl. Comp. Harm. Anal.*, 1994.

[24] O. Ersoy, editor. *Fourier Realated Transfroms, Fast Algorithms and Applications*. Prentice Hall, New Jersey, 1997.

[25] R. O. Gandulfo. Séries de Fourier e convergência. *Matemática Universitária*, (11):27–52, 1990.

[26] C. Gasquet and P. Witomski. *Fourier Analysis and Applications*. Springer-Verlag, New York, 1998.

[27] J. Geronimo, D. Hardin, and P. Massopust. Fractal functions and wavelet expansions based on several scaling functions. *J. Approx. Theory*, 78:373–401, 1994.

[28] A. Haar. Zur theorie der orthogonalen Funktionen Systeme. *Mat. Ann.*, 69:331–371, 1910.

[29] P. Halmos. *Mesaure Theory*. Van Nostrand, Princeton, N.J., 1950.

[30] F. J. Harris. On the use of windows for Harmonic Analysis with the Discrete Fourier Transform. *Proc. IEEE*, 66:51–83, 1978.

[31] R. V. Hartley. A more simmetrical Fourier analysis applied to transmission problems. In *Proceedings of the IRE*, volume 30, pages 142–150, 1942.

[32] M. T. Heideman, D. H. Johnson, and C. S. Burrus. Gauss and the history of the Fast Fourier Transform. *IEEE ASSP Magazine*, pages 14–21, 1984.

[33] E. I. Jury. *Theory and Applications of the z-Transfrm*. John Wiley & Sons, New York, 1964.

[34] P. G. Lemarié. Une nouvelle base d' ondelettes de $L^2(\mathbb{R}^n)$. *J. de Mat. Pures et Appl.*, 67:227–236, 1988.

[35] H. Loomis. *An Introduction to Abstact Harmonic Analysis*. Van Nostrand, New York, 1953.

[36] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, New York, 1998.

[37] Y. Meyer. Principe d'incertitude, bases Hilbertiennes et algèbres d'opérateurs. *Séminaire Bourbaki*, 662:1–15, 1985-1986.

[38] A. V. Oppenheim and R. W. Shafer. *Discrete-Time Signal Processing*. Prentice Hall, Englewood Cliffs, New Jersey, 1989.

[39] W. H. Press, B. P. Flannery, S. A. Teukolski, and W. T. Vetterling. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 2 edition, 1992. There are editions of this book with computer programs in Fortran, Pascal or C.

[40] W. Rudin. *Fourier Analysis on Groups*. John Wiley & Sons, New York, 1990.

[41] H. V. Sorensen, C. S. Burrus, and M. T. Heideman. *Fast Fourier Transforms Database*. PWS Publishing, Boston, 1995.

[42] J. O. Strömberg. A modified Franklin system and higher order spline systems on $\mathbb{R}^n$ as unconditional bases for Hardy spaces. In W. Beckner, A. P. Calderón, R. Fefferman, and P. W. Jones, editors, *Conf. in honor of Antoni Zygmund*, volume II, pages 475–493, New York, 1981. Wadsworth Math. Series.

[43] W. Sweldens. The lifting scheme: a new philosophy in biorthogonal wavelet constructions. In *Wavelet Applications in Signal and Image Processing III*, volume 2569, pages 68–79. SPIE, 1995.

[44] W. Sweldens. The lifting scheme: a custom-design construction of biorthogonal wavelets. *Appl. Comp. Harmon. Anal*, 3(2):186–200, 1996.

[45] W. Sweldens. The lifting scheme: a construction of second generation wavelets. *SIAM J. Math. Anal*, 29(2):51–546, 1997.

[46] B. Torrésani. Time-frequency representations: Wavelet packets and optimal decomposition. *Ann. Inst. H. Poincaré, Physique Théorique*, 56:215–234, 1992.

[47] J. S. Walker. *Fast Fourier Transforms.* CRC Press, Boca Raton, 1996.

[48] M. V. Wickerahauser. Acoustic signal compression with wavelet-packets. In C. K. Chui, editor, *Wavelts: A Tutorial of Theory and Applications*, volume 2 of *Wavelet Analysis and Its Applications*, pages 679–700. Academic Press, Boston, 1992.

[49] M. V. Wickerhauser. *Adapted Wavelet Analysis from Theory to Software.* A K Peters, Massachusets, 1994.

[50] H. Wilbraham. *Camb. and Dublin Math. J.*, 3:198, 1848.