Mathematics of the Discrete Fourier Transform (DFT)

Julius O. Smith III (jos@ccrma.stanford.edu)
Center for Computer Research in Music and Acoustics (CCRMA)
Department of Music, Stanford University
Stanford, California 94305

March 15, 2002

Contents

Preface

This reader is an outgrowth of my course entitled "Introduction to Digital Signal Processing and the Discrete Fourier Transform (DFT)[1] which I have given at the Center for Computer Research in Music and Acoustics (CCRMA) every year for the past 16 years. The course was created primarily as a first course in digital signal processing for entering Music Ph.D. students. As a result, the only prerequisite is a good high-school math background. Calculus exposure is desirable, but not required.

Outline

Below is an overview of the chapters.

- **Introduction to the DFT**
  This chapter introduces the Discrete Fourier Transform (DFT) and points out the elements which will be discussed in this reader.

- **Introduction to Complex Numbers**
  This chapter provides an introduction to complex numbers, factoring polynomials, the quadratic formula, the complex plane, Euler's formula, and an overview of numerical facilities for complex numbers in Matlab and Mathematica.

- **Proof of Euler's Identity**
  This chapter outlines the proof of Euler's Identity, which is an important tool for working with complex numbers. It is one of the critical elements of the DFT definition that we need to understand.

- **Logarithms, Decibels, and Number Systems**
  This chapter discusses logarithms (real and complex), decibels, and

[1] http://www-ccrma.stanford.edu/CCRMA/Courses/320/

number systems such as binary integer fixed-point, fractional fixed-point, one's complement, two's complement, logarithmic fixed-point, μ-law, and floating-point number formats.

- **Sinusoids and Exponentials**
  This chapter provides an introduction to sinusoids, exponentials, complex sinusoids, t_{60}, in-phase and quadrature sinusoidal components, the analytic signal, positive and negative frequencies, constructive and destructive interference, invariance of sinusoidal frequency in linear time-invariant systems, circular motion as the vector sum of in-phase and quadrature sinusoidal motions, sampled sinusoids, generating sampled sinusoids from powers of z, and plot examples using Mathematica.

- **The Discrete Fourier Transform (DFT) Derived**
  This chapter derives the Discrete Fourier Transform (DFT) as a projection of a length N signal $x(\cdot)$ onto the set of N sampled complex sinusoids generated by the N roots of unity.

- **Fourier Theorems for the DFT**
  This chapter derives various *Fourier theorems* for the case of the DFT. Included are symmetry relations, the shift theorem, convolution theorem, correlation theorem, power theorem, and theorems pertaining to interpolation and downsampling. Applications related to certain theorems are outlined, including linear time-invariant filtering, sampling rate conversion, and statistical signal processing.

- **Example Applications of the DFT**
  This chapter goes through some practical examples of FFT analysis in Matlab. The various Fourier theorems provide a "thinking vocabulary" for understanding elements of spectral analysis.

- **A Basic Tutorial on Sampling Theory**
  This appendix provides a basic tutorial on sampling theory. Aliasing due to sampling of continuous-time signals is characterized mathematically. Shannon's sampling theorem is proved. A pictorial representation of continuous-time signal reconstruction from discrete-time samples is given.

Chapter 1

Introduction to the DFT

This chapter introduces the Discrete Fourier Transform (DFT) and points out the elements which will be discussed in this reader.

1.1 DFT Definition

The *Discrete Fourier Transform (DFT)* of a signal x may be defined by

$$X(\omega_k) \triangleq \sum_{n=0}^{N-1} x(t_n) e^{-j\omega_k t_n}, \qquad k = 0, 1, 2, \ldots, N-1$$

and its *inverse* (the IDFT) is given by

$$x(t_n) = \frac{1}{N} \sum_{k=0}^{N-1} X(\omega_k) e^{j\omega_k t_n}, \qquad n = 0, 1, 2, \ldots, N-1$$

where

$$\begin{aligned}
x(t_n) &\stackrel{\Delta}{=} \text{input signal } \textit{amplitude} \text{ at time } t_n \text{ (sec)} \\
t_n &\stackrel{\Delta}{=} nT = n\text{th sampling instant (sec)} \\
n &\stackrel{\Delta}{=} \text{sample number (integer)} \\
T &\stackrel{\Delta}{=} \text{sampling period (sec)} \\
X(\omega_k) &\stackrel{\Delta}{=} \textit{Spectrum} \text{ of } x\text{, at radian frequency } \omega_k \\
\omega_k &\stackrel{\Delta}{=} k\Omega = k\text{th frequency sample (rad/sec)} \\
\Omega &\stackrel{\Delta}{=} \frac{2\pi}{NT} = \text{radian-frequency sampling interval} \\
f_s &\stackrel{\Delta}{=} 1/T = \textit{sampling rate} \text{ (samples/sec, or Hertz (Hz))} \\
N &= \text{number of samples in both time and frequency (integer)}
\end{aligned}$$

1.2 Mathematics of the DFT

In the signal processing literature, it is common to write the DFT in the more pure form obtained by setting $T = 1$ in the previous definition:

$$\begin{aligned} X(k) &\triangleq \sum_{n=0}^{N-1} x(n)e^{-j2\pi nk/N}, \qquad k = 0,1,2,\ldots,N-1 \\ x(n) &= \frac{1}{N}\sum_{k=0}^{N-1} X(k)e^{j2\pi nk/N}, \qquad n = 0,1,2,\ldots,N-1 \end{aligned}$$

where $x(n)$ denotes the input signal at time (sample) n, and $X(k)$ denotes the kth spectral sample.[1] This form is the simplest mathematically while the previous form is the easier to interpret physically.

There are two remaining symbols in the DFT that we have not yet defined:

$$\begin{aligned} j &\triangleq \sqrt{-1} \\ e &\triangleq \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = 2.71828182845905\ldots \end{aligned}$$

The first, $j = \sqrt{-1}$, is the basis for *complex numbers*. As a result, complex numbers will be the first topic we cover in this reader (but only to the extent needed to understand the DFT).

The second, $e = 2.718\ldots$, is a transcendental number defined by the above limit. In this reader, we will derive e and talk about why it comes up.

Note that not only do we have complex numbers to contend with, but we have them appearing in exponents, as in

$$s_k(n) \triangleq e^{j2\pi nk/N}$$

We will systematically develop what we mean by imaginary exponents in order that such mathematical expressions are well defined.

With e, j, and imaginary exponents understood, we can go on to prove *Euler's Identity*:

$$e^{j\theta} = \cos(\theta) + j\sin(\theta)$$

[1] Note that the definition of $x()$ has changed unless the sampling rate f_s really is 1, and the definition of $X()$ has changed no matter what the sampling rate is, since when $T = 1$, $\omega_k = 2\pi k/N$, not k.

Euler's Identity is the key to understanding the meaning of expressions like

$$s_k(t_n) \stackrel{\Delta}{=} e^{j\omega_k t_n} = \cos(\omega_k t_n) + j\sin(\omega_k t_n)$$

We'll see that such an expression defines a *sampled complex sinusoid*, and we'll talk about sinusoids in some detail, from an audio perspective.

Finally, we need to understand what the summation over n is doing in the definition of the DFT. We'll learn that it should be seen as the computation of the *inner product* of the signals x and s_k, so that we may write the DFT using inner-product notation as

$$X(k) \stackrel{\Delta}{=} \langle x, s_k \rangle$$

where

$$s_k(n) \stackrel{\Delta}{=} e^{j2\pi nk/N}$$

is the sampled complex sinusoid at (normalized) radian frequency $\omega_k = 2\pi k/N$, and the inner product operation is defined by

$$\langle x, y \rangle \stackrel{\Delta}{=} \sum_{n=0}^{N-1} x(n)\overline{y(n)}$$

We will show that the inner product of x with the kth "basis sinusoid" s_k is a measure of "how much" of s_k is present in x and at "what phase" (since it is a complex number).

After the foregoing, the inverse DFT can be understood as the *weighted sum of projections* of x onto $\{s_k\}_{k=0}^{N-1}$, i.e.,

$$x(n) \stackrel{\Delta}{=} \sum_{k=0}^{N-1} \tilde{X}_k s_k(n)$$

where

$$\tilde{X}_k \stackrel{\Delta}{=} \frac{X(k)}{N}$$

is the (actual) *coefficient of projection* of x onto s_k. Referring to the whole signal as $x \stackrel{\Delta}{=} x(\cdot)$, the IDFT can be written as

$$x \stackrel{\Delta}{=} \sum_{k=0}^{N-1} \tilde{X}_k s_k$$

Note that both the *basis sinusoids* s_k and their coefficients of projection $\tilde{X}_k$ are *complex*.

Having completely understood the DFT and its inverse mathematically, we go on to proving various *Fourier Theorems*, such as the "shift theorem," the "convolution theorem," and "Parseval's theorem." The Fourier theorems provide a basic thinking vocabulary for working with signals in the time and frequency domains. They can be used to answer questions like

> What happens in the frequency domain if I do [x] in the time domain?

Finally, we will study a variety of practical spectrum analysis examples, using primarily Matlab to analyze and display signals and their spectra.

1.3 DFT Math Outline

In summary, understanding the DFT takes us through the following topics:

1. Complex numbers
2. Complex exponents
3. Why e?
4. Euler's formula
5. Projecting signals onto signals via the inner product
6. The DFT as the coefficient of projection of a signal x onto a sinusoid
7. The IDFT as a weighted sum of sinusoidal projections
8. Various Fourier theorems
9. Elementary time-frequency pairs
10. Practical spectrum analysis in matlab

We will additionally discuss practical aspects of working with sinusoids, such as *decibels* (dB) and display techniques.

Chapter 2

Complex Numbers

This chapter provides an introduction to complex numbers, factoring polynomials, the quadratic formula, the complex plane, Euler's formula, and an overview of numerical facilities for complex numbers in Matlab and Mathematica.

2.1 Factoring a Polynomial

Remember "factoring polynomials"? Consider the second-order polynomial

$$p(x) = x^2 - 5x + 6$$

It is second-order because the highest power of x is 2 (only non-negative integer powers of x are allowed in this context). The polynomial is also *monic* because its leading coefficient, the coefficient of x^2, is 1. Since it is second order, there are at most two real *roots* (or *zeros*) of the polynomial. Suppose they are denoted x_1 and x_2. Then we have $p(x_1) = 0$ and $p(x_2) = 0$, and we can write

$$p(x) = (x - x_1)(x - x_2)$$

This is the *factored form* of the monic polynomial $p(x)$. (For a non-monic polynomial, we may simply divide all coefficients by the first to make it monic, and this doesn't affect the zeros.) Multiplying out the symbolic factored form gives

$$p(x) = (x - x_1)(x - x_2) = x^2 - (x_1 + x_2)x + x_1 x_2$$

Comparing with the original polynomial, we find we must have

$$\begin{aligned} x_1 + x_2 &= 5 \\ x_1 x_2 &= 6 \end{aligned}$$

This is a system of two equations in two unknowns. Unfortunately, it is a *nonlinear* system of two equations in two unknowns.[1] Nevertheless, because it is so small, the equations are easily solved. In beginning algebra, we did them by hand. However, nowadays we can use a computer program such as Mathematica:

```
In[]:=
        Solve[{x1+x2==5, x1 x2 == 6}, {x1,x2}]
Out[]:
        {{x1 -> 2, x2 -> 3}, {x1 -> 3, x2 -> 2}}
```

Note that the two lists of substitutions point out that it doesn't matter which root is 2 and which is 3. In summary, the factored form of this simple example is

$$p(x) = x^2 - 5x + 6 = (x - x_1)(x - x_2) = (x - 2)(x - 3)$$

Note that polynomial factorization rewrites a monic nth-order polynomial as the product of n *first-order* monic polynomials, each of which contributes one zero (root) to the product. This factoring business is often used when working with *digital filters*.

2.2 The Quadratic Formula

The general second-order polynomial is

$$p(x) \stackrel{\Delta}{=} ax^2 + bx + c$$

where the coefficients a, b, c are any real numbers, and we assume $a \neq 0$ since otherwise it would not be second order. Some experiments plotting

[1] "Linear" in this context means that the unknowns are multiplied only by constants—they may not be multiplied by each other or raised to any power other than 1 (e.g., not squared or cubed or raised to the 1/5 power). Linear systems of N equations in N unknowns are very easy to solve compared to *nonlinear* systems of N equations in N unknowns. For example, Matlab or Mathematica can easily handle them. You learn all about this in a course on *Linear Algebra* which is highly recommended for anyone interested in getting involved with signal processing. Linear algebra also teaches you all about *matrices* which we will introduce only briefly in this reader.

$p(x)$ for different values of the coefficients leads one to guess that the curve is always a scaled and translated *parabola*. The canonical parabola centered at $x = x_0$ is given by

$$y(x) = d(x - x_0)^2 + e$$

where d determines the width and e provides an arbitrary vertical offset. If we can find d, e, x_0 in terms of a, b, c for any quadratic polynomial, then we can easily factor the polynomial. This is called "completing the square." Multiplying out $y(x)$, we get

$$y(x) = d(x - x_0)^2 + e = dx^2 - 2dx_0x + dx_0^2 + e$$

Equating coefficients of like powers of x gives

$$\begin{aligned} d &= a \\ -2dx_0 &= b \quad \Rightarrow \quad x_0 = -b/(2a) \\ dx_0^2 + e &= c \quad \Rightarrow \quad e = c - b^2/(4a) \end{aligned}$$

Using these answers, any second-order polynomial $p(x) = ax^2 + bx + c$ can be rewritten as a scaled, translated parabola

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

In this form, the roots are easily found by solving $p(x) = 0$ to get

$$\boxed{x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}}$$

This is the general *quadratic formula*. It was obtained by simple algebraic manipulation of the original polynomial. There is only one "catch." What happens when $b^2 - 4ac$ is negative? This introduces the square root of a negative number which we could insist "does not exist." Alternatively, we could invent complex numbers to accommodate it.

2.3 Complex Roots

As a simple example, let $a = 1$, $b = 0$, and $c = 4$, i.e.,

$$p(x) = x^2 + 4$$

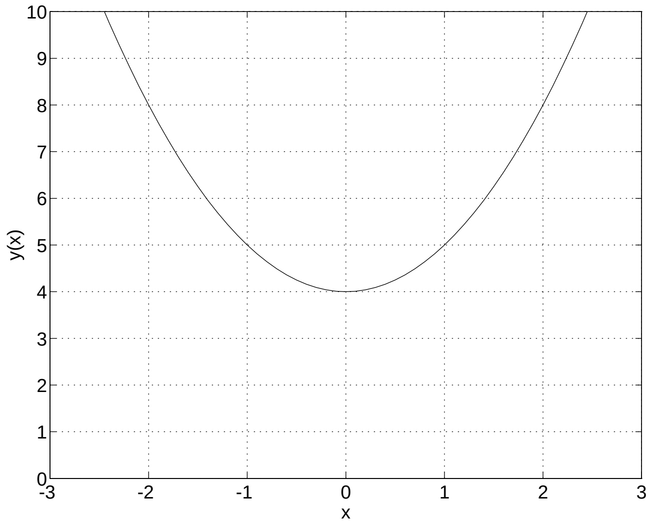

Figure 2.1: An example parabola defined by $p(x) = x^2 + 4$.

As shown in Fig. 2.1, this is a parabola centered at $x = 0$ (where $p(0) = 4$) and reaching upward to positive infinity, never going below 4. It has no zeros. On the other hand, the quadratic formula says that the "roots" are given formally by $x = \pm\sqrt{-4} = \pm 2\sqrt{-1}$. The square root of any negative number $c < 0$ can be expressed as $\sqrt{|c|}\sqrt{-1}$, so the only new algebraic object is $\sqrt{-1}$. Let's give it a name:

$$\boxed{j \stackrel{\Delta}{=} \sqrt{-1}}$$

Then, formally, the roots of of x^2+4 are $\pm 2j$, and we can formally express the polynomial in terms of its roots as

$$p(x) = (x + 2j)(x - 2j)$$

We can think of these as "imaginary roots" in the sense that square roots of negative numbers don't really exist, or we can extend the concept of "roots" to allow for *complex numbers*, that is, numbers of the form

$$z = x + jy$$

where x and y are real numbers, and $j^2 \stackrel{\Delta}{=} -1$.

It can be checked that all algebraic operations for real numbers[2] apply equally well to complex numers. Both real numbers and complex numbers are examples of a mathematical *field*. Fields are *closed* with respect to multiplication and addition, and all the rules of algebra we use in manipulating polynomials with real coefficients (and roots) carry over unchanged to polynomials with complex coefficients and roots. In fact, the rules of algebra become simpler for complex numbers because, as discussed in the next section, we can *always* factor polynomials completely over the field of complex numbers while we cannot do this over the reals (as we saw in the example $p(x) = x^2 + 4$).

2.4 Fundamental Theorem of Algebra

Every nth-order polynomial possesses exactly n complex roots.

This is a very powerful algebraic tool. It says that given any polynomial

$$\begin{aligned} p(x) &= a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \cdots + a_2 x^2 + a_1 x + a_0 \\ &\triangleq \sum_{i=0}^{n} a_i x^i \end{aligned}$$

we can *always* rewrite it as

$$\begin{aligned} p(x) &= a_n (x - z_n)(x - z_{n-1})(x - z_{n-2}) \cdots (x - z_2)(x - z_1) \\ &\triangleq a_n \prod_{i=1}^{n} (x - z_i) \end{aligned}$$

where the points z_i are the polynomial roots, and they may be real or complex.

2.5 Complex Basics

This section introduces various notation and terms associated with complex numbers. As discussed above, complex numbers are devised by introducing the square-root of -1 as a primitive new algebraic object among

[2] multiplication, addition, division, distributivity of multiplication over addition, commutativity of multiplication and addition.

real numbers and manipulating it symbolically as if it were a real number itself:

$$\boxed{j \stackrel{\Delta}{=} \sqrt{-1}}$$

Mathemeticians and physicists often use i instead of j as $\sqrt{-1}$. The use of j is common in engineering where i is more often used for electrical current.

As mentioned above, for any negative number $c < 0$, we have $\sqrt{c} = \sqrt{(-1)(-c)} = j\sqrt{|c|}$, where $|c|$ denotes the absolute value of c. Thus, every square root of a negative number can be expressed as j times the square root of a positive number.

By definition, we have

$$\begin{aligned}
j^0 &= 1 \\
j^1 &= j \\
j^2 &= -1 \\
j^3 &= -j \\
j^4 &= 1 \\
&\cdots
\end{aligned}$$

and so on. Thus, the sequence $x(n) \stackrel{\Delta}{=} j^n$, $n = 0, 1, 2, \ldots$ is a periodic sequence with period 4, since $j^{n+4} = j^n j^4 = j^n$. (We'll learn later that the sequence j^n is a sampled complex sinusoid having frequency equal to one fourth the sampling rate.)

Every *complex number* z can be written as

$$\boxed{z = x + jy}$$

where x and y are real numbers. We call x the *real part* and y the *imaginary part*. We may also use the notation

$$\begin{aligned}
\text{re}\{z\} &= x \qquad (\text{“the real part of } z = x + jy \text{ is } x\text{”}) \\
\text{im}\{z\} &= y \qquad (\text{“the imaginary part of } z = x + jy \text{ is } y\text{”})
\end{aligned}$$

Note that the real numbers are the subset of the complex numbers having a zero imaginary part ($y = 0$).

The rule for *complex multiplication* follows directly from the definition of the imaginary unit j:

$$\begin{aligned}
z_1 z_2 &\stackrel{\Delta}{=} (x_1 + jy_1)(x_2 + jy_2) \\
&= x_1 x_2 + jx_1 y_2 + jy_1 x_2 + j^2 y_1 y_2 \\
&= (x_1 x_2 - y_1 y_2) + j(x_1 y_2 + y_1 x_2)
\end{aligned}$$

In some mathematics texts, complex numbers z are defined as ordered pairs of real numbers (x, y), and algebraic operations such as multiplication are defined more formally as operations on ordered pairs, e.g., $(x_1, y_1) \cdot (x_2, y_2) \stackrel{\Delta}{=} (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2)$. However, such formality tends to obscure the underlying simplicity of complex numbers as a straightforward extension of real numbers to include $j \stackrel{\Delta}{=} \sqrt{-1}$.

It is important to realize that complex numbers can be treated algebraically just like real numbers. That is, they can be added, subtracted, multiplied, divided, etc., using exactly the same rules of algebra (since both real and complex numbers are mathematical *fields*). It is often preferable to think of complex numbers as being the true and proper setting for algebraic operations, with real numbers being the limited subset for which $y = 0$.

To explore further the magical world of complex variables, see any textbook such as [1, 2].

2.5.1 The Complex Plane

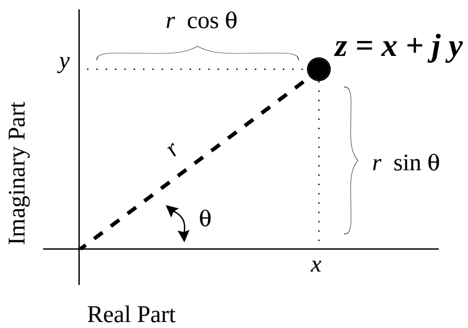

Figure 2.2: Plotting a complex number as a point in the complex plane.

We can plot any complex number $z = x + jy$ in a plane as an ordered pair (x, y), as shown in Fig. 2.2. A *complex plane* is any 2D graph in which the horizontal axis is the *real part* and the vertical axis is the *imaginary part* of a complex number or function. As an example, the number j has coordinates $(0, 1)$ in the complex plane while the number 1 has coordinates $(1, 0)$.

Plotting $z = x + jy$ as the point (x, y) in the complex plane can be viewed as a plot in *Cartesian* or *rectilinear* coordinates. We can also express complex numbers in terms of *polar coordinates* as an ordered pair (r, θ), where r is the distance from the origin $(0, 0)$ to the number being plotted, and θ is the angle of the number relative to the positive real coordinate axis (the line defined by $y = 0$ and $x > 0$). (See Fig. 2.2.)

Using elementary *geometry*, it is quick to show that conversion from rectangular to polar coordinates is accomplished by the formulas

$$\boxed{\begin{array}{rcl} r & = & \sqrt{x^2 + y^2} \\ \theta & = & \tan^{-1}(y/x). \end{array}}$$

The first equation follows immediately from the *Pythagorean theorem* , while the second follows immediately from the definition of the *tangent* function. Similarly, conversion from polar to rectangular coordinates is simply

$$\boxed{\begin{array}{rcl} x & = & r\ \cos(\theta) \\ y & = & r\ \sin(\theta). \end{array}}$$

These follow immediately from the definitions of cosine and sine, respectively,

2.5.2 More Notation and Terminology

It's already been mentioned that the rectilinear coordinates of a complex number $z = x + jy$ in the complex plane are called the *real part* and *imaginary part*, respectively.

We also have special notation and various names for the radius and angle of a complex number z expressed in polar coordinates (r, θ):

$$\begin{array}{rcl} r & \stackrel{\Delta}{=} & |z| = \sqrt{x^2 + y^2} \\ & = & \textit{modulus}, \textit{magnitude}, \textit{absolute value}, \textit{norm}, \text{ or } \textit{radius} \text{ of } z \\ \theta & \stackrel{\Delta}{=} & \angle z = \tan^{-1}(y/x) \\ & = & \textit{angle}, \textit{argument}, \text{ or } \textit{phase} \text{ of } z \end{array}$$

The *complex conjugate* of z is denoted $\overline{z}$ and is defined by

$$\overline{z} \stackrel{\Delta}{=} x - jy$$