Presentation of article "Tom, Dick and Mary discover the DFT" by J.R. Deller, Jr.

Ruben Spaans

October 27, 2009

Motivation for the article

The author complains over the fact that students learn discrete Fourier transform (hereby referred to as DFT) without learning about continuous-time Fourier analysis.

The article follows three students who discover the DFT, having only knowledge about analog methods.

Non-rigorous definitions

Fourier transform The Fourier transform (FT) is an operation that transforms a function $f(t)$ from the time (or spatial) domain into a function $F(f)$ in the frequency domain.

Inverse Fourier transform The inverse Fourier transform is the reverse of the above; a transformation of a function $F(f)$ from the frequency domain to $f(t)$ in the time (or spatial) domain.

Fourier series A Fourier series (FS) of a periodic function f is the decomposition of f into a sum of sine/cosine functions or complex exponentials.

Homework

Tom, Dick and Mary are working on an assignment for a course in signal and linear system analysis. The assignment involves, among other things, plotting of magnitude and phase spectra based on the Fourier transform.

The first signal they encountered was $x_1(t)$ and its Fourier integral:

$$X_1(f) = \int_{-\infty}^{\infty} x_1(t) e^{-j2\pi f t}\, dt. \tag{1}$$

(Please note the usage of j for $\sqrt{-1}$. Also note that the author uses a different definition of Fourier transform that what we learned in Matematikk 4, by omitting the scalar $\frac{1}{\sqrt{2\pi}}$.) Their idea was to take closely spaced samples, and draw the line between them to achieve something that looked like a continuous plot. However, they didn't know how to sample from X_1 using discrete computation. They couldn't see how to compute the integral.

However, they soon realise that the Fourier series (FS) resulted in a sort of sampled frequency domain. Even though $x_1(t)$ is not a periodic signal, only one period is needed in constructing a FS.

They made $x_1(t)$ periodic by choosing a general period T_y such that T_y is not smaller than the width of $x_1(t)$ (to avoid overlapping), and the resulting periodic signal $y(t)$.

Then,

$$y(t) = \sum_{k=-\infty}^{\infty} x_1(t - kT_y). \tag{2}$$

Then they proceeded by writing the Fourier series for $y(t)$, given by

$$y(t) = \sum_{m=-\infty}^{\infty} \alpha_m e^{j2\pi m f_y t}, \;\; f_y = \frac{1}{T_y} \tag{3}$$

$$\alpha_m = \int_{-T_y/2}^{T_y/2} y(t) e^{-j2\pi m f_y t}\, dt \quad m \in \mathbb{Z}. \tag{4}$$

(This is, again, a different form than the definition from Matematikk 4.) They noticed that the evaluation of α_m only used one period, so they could just replace $y(t)$ by $x_1(t)$, the original signal.

Then, they saw that the coefficients α_m could be written in terms of the Fourier transform of $x_1(t)$ (with a scale factor):

$$\alpha_m = \frac{1}{T_y} X_1(m f_y) \quad m \in \mathbb{Z}. \tag{5}$$

(For details of the calculations, see the original article.)

They had FS coefficients for a periodic version of $x_1(t)$ whose copies doesn't overlap.

What happens if they choose $f_y = 1/T_y$ where $T_y < T_1$ (where T_1 is the width of $x_1(t)$), that is, $y(t)$ has overlapping?

They concluded that what they had the overlapping case. But then, they were uncertain whether this manufactured signal was $y(t)$ (they called it $y'(t)$), and went on to take the FT of the FS they have constructed. They did some strange calculations involving the dirac delta function, and manage to conclude that $y'(t) = y(t)$.

To get one period of $y(t)$ to be exactly $x_1(t)$, they had to make sure that frequencies of the samples of $X_1(f)$ at $mf_y = m/T_y$ are close enough together: $T_y > T_1$. Still, they had to get rid of the continuous signals in both time and frequency domains in order to be able to do plots.

They didn't know what $X_1(f)$ looked like.

They thought that it should be possible to reverse the process they had just used in order to get samples in the time domain: It should be possible to compute $x_1(t)$ if they knew samples of $X_1(f)$.

So they considered the periodic function

$$Y(f) = \sum_{n=-\infty}^{\infty} X_1(f - nf_s). \tag{6}$$

They speculated that this function could be represented by a "Fourier series" were related to samples of time. So they proceed by computing the "Fourier transform" of $X_1(f)$, calling it $x_1'(t)$:

$$x_1'(t) = \int_{-\infty}^{\infty} X_1(f)e^{-j2\pi tf}\, df \tag{7}$$

which was very close to the time signal

$$x_1(t) = \int_{-\infty}^{\infty} X_1(f)e^{j2\pi tf}\, df = x_1'(-t). \tag{8}$$

(Actually, an actual inverse Fourier transform will be the "mirror function" of $x_1'(t)$.) Since $Y(f)$ was periodic, they constructed the "FS", letting $T_s = 1/f_s$:

$$Y(f) = \sum_{n=-\infty}^{\infty} \beta_n e^{j2\pi nT_s f}, \quad T_s = 1/f_s \tag{9}$$

$$\beta_n = \frac{1}{f_s}\int_{-f_s/2}^{f_s/2} Y(f)e^{-j2\pi nT_s f}\, df = \frac{1}{f_s}x_1'(nT_s) = \frac{1}{f_s}x_1'(-nT_s). \tag{10}$$

The "FS" coefficients are scaled samples of the reversed time waveform $x_1'(t) = x_1(-t)$. Rewriting $Y(f)$ in terms of $x_1(t)$ yields

$$Y(f) = \sum_{n=-\infty}^{\infty} \frac{1}{f_s}x_1(-nT_s)e^{j2\pi nT_s f} \tag{11}$$

They didn't like the idea of using samples in reversed time, so they replaced $-n$ by n:

$$Y(f) = \frac{1}{f_s}\sum_{n=-\infty}^{\infty} x_1(nT_s)e^{-j2\pi nT_s f} \tag{12}$$

They discovered that as long as they chose $f_s = 1/T_s > 2f_b$[1], then the first period of $Y(f)$ would be exactly $X_1(f)$. Thus they could calculate $Y(f) =$

[1] With this, they have actually rediscovered Shannon's sampling theorem, which says that a complete continuous-time signal $x_1(t)$ is recoverable from the time samples $x_1(nT_s)$ as long as the chosen sample rate is high enough, ie. $f_s = 1/T_s > 2f_b$ where f_b is the bandwidth of the signal.

$X_1(f)$ at any frequency they wanted using equation 12, by using discrete samples of $x_1(t)$ for $t = \ldots, -2T_s, T_s, 0, T_s, 2T_s, \ldots$. For any frequency f_0 of which they wanted to calculate $Y(f_0)$, they had to ensure that $-f_b \leq f_0 \leq f_b$ (the range over which $Y(f) = X_1(f)$) and $f_s \geq 2f_b$. The bandwidth f_b was unknown, so they had to guess the sample period T_s to use in each case, this sometimes gave bad results (see next paragraph for more about this). (The article doesn't mention how they evaluated the infinite sum, but as $|n|$ grows, one will eventually sample "outside" of $x_1(t)$, hence it suffices to evaluate a finite number of terms.)

Some notes about sampling from a signal (which the professor told the students). As long as $x_1(t)$ is sampled fast enough, the "FS" $Y(f)$ is an excellent candidate for spectral analysis of the original signal, since it represents an exactly periodic version of $X_1(f)$. But when samples are not taken fast enough, $Y(f)$ will represent a periodic, but *aliased* version of $X_1(f)$. The term *aliasing* refers to the overlap in the frequency spectrum which causes frequencies above $f_s/2$ in the signal to be confused with those in the band $0 \leq f < f_s/2$. This explains why the students got some unsatisfactory plots. They had chosen an arbitrary sample rate, which, in the case of signal $x_2(t)$, wasn't high enough to encompass its bandwidth. See the article, page 42 for more detailed information.

The professor also told the students that their result, the FS of $Y(f)$ that they discovered, was actually very close to the *discrete-time Fourier transform* (DTFT):

$$Y_{\text{DTFT}}(f) = \sum_{n=-\infty}^{\infty} x_1(nT_s)e^{-j2\pi f nT_s}. \tag{13}$$

Furthermore, the professor told them that the DTFT allows people to compute a periodic replica of the continuous FT using just time samples. This enables people to perform meaningful spectral analysis using a computer.

Onward to the DFT

The professor gave the students a new challenge (which, by the way, they weren't very happy about): To find a transform to convert back and forth between discrete sets of samples, between the time and frequency domains. They could assume that their next signal, $x_3(t)$, is finite in time, and that they could take N samples, $x_3(0), x_3(T_s), x_3(2T_s), \ldots, x_3((N-1)T_s)$.

After some pondering and the realisation that they have only N time samples, equation 12 could be simplified to

$$Y(f) = \frac{1}{f_s} \sum_{n=0}^{N-1} x_3(nT_s) e^{-j2\pi T_s f} \tag{14}$$

where $Y(f)$ is the periodic version of $X_3(f)$. The expression for $y(t)$, using samples at spacing $f_s = 1/NT_s = f_s/N$ now becomes

$$y(t) = \sum_{k=-\infty}^{\infty} x_3(t - kNT_s), \tag{15}$$

and taking samples at times nT_s for $n = 0, 1, \ldots, N-1$:

$$y(nT_s) = \sum_{k=-\infty}^{\infty} x_3(nT_s - kNT_s). \tag{16}$$

They had a hunch that taking samples from $Y(f)$ with the correct spacing would give back a function that matched $y(t)$ at the points where $y(t)$ was originally sampled.

They decided that the periodicity of $Y(f)$ was throwing them off, so they defined $Y_3(f)$ which was equal to $Y(f)$ for $\frac{-f_s}{2} \leq f \leq \frac{f_s}{2}$ and 0 elsewhere, forcing $y_3(t)$ to be a continuous-time signal which would be equal to the original $x_3(t)$ if $x_3(t)$ had been bandlimited and sampled fast enough. However, because of potential aliasing, some uncertainty ensued about whether samples from $y_3(t)$ would be equal to samples of $x_3(t)$. They decided to proceed anyway.

They created a periodic extension of $y_3(t)$:

$$w(t) = \sum_{i=-\infty}^{\infty} y_3(t - iT_w) \tag{17}$$

where $f_w = 1/T_w$ was the sample interval of $Y_3(f)$ they would use, where $T_w = MT_s$ for some positive integer M.

They proceeded to take the FS:

$$w(t) = \sum_{k=-\infty}^{\infty} \gamma_k e^{j2\pi k f_w t} \tag{18}$$

$$\gamma_k = \frac{1}{T_w} \int_{-T_w/2}^{T_w/2} w(t) e^{-j2\pi k f_w t} \tag{19}$$

From past experience, they knew that the FS coefficients could be written in terms of the FT of the nonperiodic version of the function:

$$\gamma_k = \frac{1}{T_w} Y_3(k f_w), \tag{20}$$

and hence

$$w(t) = \frac{1}{T_w} \sum_{k=-\infty}^{\infty} Y_3(kf_w) e^{j2\pi k f_w t}. \tag{21}$$

They noticed that they could replace $Y_3(f)$ by $Y(f)$:

$$w(t) = \frac{1}{T_w} \sum_{k} Y(kf_w) e^{j2\pi k f_w t} \tag{22}$$

by summing over all k satisfying $\frac{-f_s}{2} \leq kf_w \leq \frac{f_s}{2}$.

In order to not be too confused, we recall that:

- $x_3(t)$ is the original function, and $y(t)$ is the periodic version of $x_3(t)$.
- $Y(f)$ is the FT of $y(f)$, and $Y_3(f)$ is the version of $Y(f)$ which is 0 when $f < \frac{-f_s}{2}$ or $f > \frac{f_s}{2}$.
- $y_3(t)$ is the inverse FT of $Y_3(f)$, and $w(t)$ is the periodic version of $y_3(t)$. The students' hope is that $w(t) = x_3(t)$ for all the sampled points $t = nT_s$ for integers n.

They evaluated $w(t)$ as the discrete times which were multiples of T_s:

$$w(nT_s) = \frac{1}{T_w} \sum_{k=-\infty}^{\infty} Y(kf_w) e^{j2\pi f_w n T_s}. \tag{23}$$

They discovered that $y_3(nT_s)$ can be related to β_n, the FS coefficients of $Y(f)$:

$$\beta_n = \frac{1}{f_s} \int_{-f_s/2}^{f_s/2} Y(f) e^{-j2\pi n T_s f}\, df = \frac{1}{f_s} x_3(nT_s) = \frac{1}{f_s} x_3(nT_s). \tag{24}$$

So

$$y_3(nT_s) = \int_{-f_s/2}^{f_s/2} Y_3(f) e^{j2\pi f n T_s}\, df \tag{25}$$

$$= \int_{-f_s/2}^{f_s/2} Y(f) e^{j2\pi f n T_s}\, df \tag{26}$$

$$= f_s \beta_{-n} = x_3(nT_s). \tag{27}$$

Equation 26 follows from 25 because $Y_3(f) = Y(f)$ for all f in the integral interval. As we recall, $w(t)$ was a periodic version of $y_3(t)$ with spacing $T_w = MT_s$. $y_3(t) = x_3(t)$ when t is a multiple of T_s, so

$$w(nT_s) = \sum_{i=-\infty}^{\infty} x_3(nT_s - iMT_s). \tag{28}$$

Hence, $w(nT_s)$ was a periodic, possibly aliased version of $x_3(nT_s)$. So, as long as $M \geq N$, the samples $w(0), w(T_s), \ldots, w((N-1)T_s)$ would be the ones they were looking for and could be calculated from samples of $Y(f)$, which in turn could be calculated from $x_3(t)$ which is equal to $w(t)$. It means they found their discrete transform!

Setting $M = N$, and recalling that $T_w = MT_s$ and $f_w = \frac{1}{MT_s}$ and equation 22

$$x_3(nT_s) = w(NT_s) = \frac{1}{NT_s} \sum_k Y\left(\frac{kf_s}{N}\right) e^{j2\pi k f_s n T_s/N} \tag{29}$$

$$= \frac{1}{NT_s} \sum_k Y\left(\frac{kf_s}{N}\right) e^{j2\pi kn/N} \tag{30}$$

for $n = 0, 1, \ldots, N-1$ and k such that $\frac{-f_s}{2} \leq \frac{kf_s}{N} \leq \frac{f_s}{2}$.

For the other way around, they recalled equation 12. Using $x_3(t)$ instead and remembering that $x_3(nT_s) = 0$ except for $n = 0, 1, \ldots, N-1$, they got

$$Y\left(\frac{kf_s}{N}\right) \sum_{n=0}^{N-1} x_3(nT_s) e^{-j2\pi nk/N} \tag{31}$$

for k satisfying $\frac{-f_s}{2} \leq \frac{kf_s}{N} \leq \frac{f_s}{2}$. They proceeded to produce some more plots and presented their results to the professor.

Their result was close to the DTFT and the professor suggested they compute samples from the DTFT instead, which only differ from equation 31 by a scale factor:

$$Y_{\text{DTFT}}\left(\frac{kf_s}{N}\right) = \sum_{n=0}^{N-1} x_3(nT_s) e^{-j2\pi nk/N}. \tag{32}$$

For the inverse DTFT:

$$x_3(nT_s) = \frac{1}{N} \sum_k Y_{\text{DTFT}}\left(\frac{kf_s}{N}\right) e^{j2\pi k/N} \tag{33}$$

for $n = 0, 1, \ldots, N-1$ and k such that $\frac{-f_s}{2} \leq \frac{kf_s}{N} \leq \frac{f_s}{2}$.

There is an inconvenience with the condition for k above. It can be slightly simplified to $k = -\frac{N-1}{2}, \ldots, 0, \ldots, \frac{N-1}{2}$ if N is odd and $k = -\frac{N}{2}, \ldots, 0, \ldots, \frac{N}{2}$ if N is even. To avoid two separate formulas, use the fact that the numbers $Y_{\text{DTFT}}\left(\frac{kf_s}{N}\right) e^{j2\pi kn/N}$ are periodic with period N. It means we can compute for $k = 0, 1, \ldots, N-1$, no matter if n is odd or even.

We can then use the following inverse DTFT:

$$x_3(nT_s) = \frac{1}{N} \sum_{k=0}^{N-1} Y_{\text{DTFT}}\left(\frac{kf_s}{N}\right) e^{j2\pi kn/N} \tag{34}$$

for $n = 0, 1, \ldots, N-1$.

Equations 32 and 34 are the essence of the *discrete Fourier transform.*

It is customary to normalize the problem by reindexing the time and frequency axes so that $f_s = T_s = 1$ is used. We then get the most common form of the DFT:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi k/N}, \quad k = 0, 1, 2, \ldots, N-1 \tag{35}$$

and its inverse transform,

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j2\pi kn/N}, \quad n = 0, 1, 2, \ldots, N-1. \tag{36}$$

In the end, the students have learned the hard way to derive the DFT from the starting point of having to plot the frequency spectra of continuous functions, armed with knowledge of the continuous Fourier transform. The professor thought this was a good thing, and that they had gained a deeper understanding of the DFT and will allow them to interpret their results better.

The end

I will point out that there also exists a presentation of the same article from last year's version of this course.

Errata

As last year's presenter so kindly pointed out: There are several typos in the equations in the article. Those errors are corrected when I have used the same formulas in this presentation, but they are repeated here for convenience. (I cannot guarantee that I haven't introduced new errors in my presentation, though.)

- *Equation 4, page 38:* There should be a $\frac{1}{T_y}$ in front of the integral.
- *Equation 8, page 39:* There shold be an $=$ between $Y'(f)$ and $\frac{1}{T_y}$.
- *Equation 21, page 44:* The equation doesn't match its supposed graph (Figure 6, page 43). Something along the lines of $0.9 \cdot 1.1^{-1000t}[u(t) - u(t - 0.0151)]$ seems to match reasonably well.

- *Top right of page 44:* The sequence $x_3(0), x_3(T_s), x_3(2T_s), \ldots$ should end in $x_3((N-1)T_s)$ (the parantheses were missing).
- *Top left of page 45:* "then Eq. 7 would produce, according to Eq. 3:" - I (where "I" is last year's presenter) think he refers to 2 or 9, not 3. "And, if we used Eq. 7 at times nT_s" - probably 2, 9, or 23.
- *Equation 41, page 47:* The argument to Y_{DTFT} should be $\frac{kf_s}{N}$.

Bibliography

J.R.Deller, Jr.: Tom, Dick, and Mary Discover the DFT, IEEE signal processing magazine, April 1994.

Åsmund Eldhuset: Article review: Tom, Dick, and Mary Discover the DFT, TDT24 presentation, 2008.