A remark on zero-padding for increased frequency resolution

Fredrik Lindsten

November 4, 2010

1 Introduction

A common tool in frequency analysis of sampled signals is to use zero-padding to increase the frequency resolution of the discrete Fourier transform (DFT). By appending artificial zeros to the signal, we obtain a denser frequency grid when applying the DFT. At first this might seem counterintuitive and hard to understand. It is the purpose of this document to give a time-domain explanation of zero-padding, which hopefully can help to increase the understanding of why we obtain the results that we do.

2 Truncated DTFT

The discrete-time Fourier transform (DTFT) of a discrete-time signal x is given by

$$X_T(e^{i\omega T}) = T \sum_{k=-\infty}^{\infty} x[k] e^{-i\omega k T}. \tag{1}$$

Now, assume that we have observed/measured N values of the signal x, i.e., we know

$$x[0],\ x[1],\ \ldots,\ x[N-1], \tag{2}$$

but apart from this the signal is unknown. This, very realistic scenario, leads to an immediate difficulty when computing the DTFT, since the summation in (1) ranges over unknown values of x. To be able to cope with this problem we need to make some assumption about the behavior of the signal outside the known range.

One possibility is to assume that the unknown values are zero, as illustrated in Figure 1. If we make use of this assumption when computing the DTFT we obtain the truncated DTFT,

$$X_T^{(N)}(e^{i\omega T}) = T \sum_{k=0}^{N-1} x[k] e^{-i\omega k T}. \tag{3}$$

It is important to remember that, if the "true" signal x is non-zero outside the sampled range, then the truncated DTFT is an approximation of the "true" DTFT.

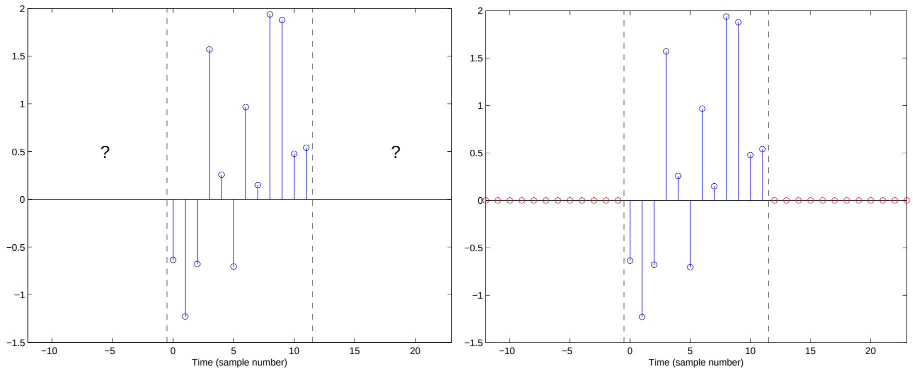

Figure 1: (Left) Sampled signal, the values outside the sampled range are unknown. (Right) One possible approximation, we assume that the signal is zero outside the sampled range.

3 DFT

The discrete Fourier transform (DFT) is often presented as a discrete approximation of the truncated DTFT. In other words, since the (truncated) DTFT is frequency continuous (it is a function of the continuous variable ω), it is problematic to represent and to work with using computers. Hence, to obtain something that is manageable, we "sample" the truncated DTFT at a discrete set of frequencies. This, frequency discrete, representation is the DFT of the signal.

However, it is possible make another interpretation of the DFT. To do so we return to the original problem, namely that we need to make some assumption about the behavior of the signal x outside the known range. Now, instead of assuming that it is zero, as we did for the truncated DTFT, we assume that the signal is periodic with period N. This assumption is illustrated in Figure 2. Hence, we claim that periodicity of the signal will turn the DTFT into the DFT. Unfortunately, to show this is not as straightforward as plugging the periodic signal into the definition of the DTFT (1) and do the computations. The reason is that a periodic signal is not of bounded energy (not in ℓ_1), meaning that its DTFT does not exist! We shall throughout this section still try to motivate the proposition that periodicity will "turn the DTFT into the DFT".

To start with, recall the definitions of the DFT and the inverse DFT, respectively,

$$X[n] = \sum_{k=0}^{N-1} x[k] e^{-\frac{2\pi i}{N} nk}, \tag{4a}$$

$$x[k] = \frac{1}{N} \sum_{n=0}^{N-1} X[n] e^{\frac{2\pi i}{N} nk}, \tag{4b}$$

$$n,\, k = 0,\, \ldots,\, N-1. \tag{4c}$$

As indicated by (4c), the definition is only valid for n and k ranging from 0 to $N-1$. However, if we relax this condition, implicit periodicity of the signal

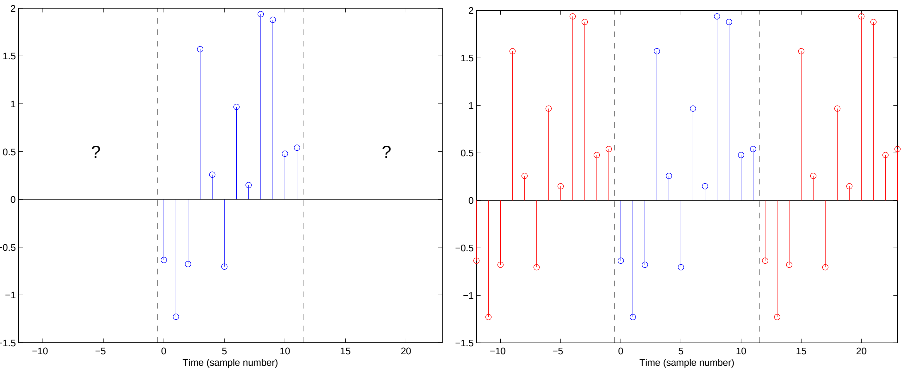

Figure 2: (Left) Sampled signal, the values outside the sampled range are unknown. (Right) One possible approximation, we assume that the signal is periodic with period N.

can be seen already in the DFT definition. More precisely, let $\bar{k}$ be any natural number and decompose it as $\bar{k} = Np + k$ for $k \in [0, N-1]$ and p a natural number. Consider,

$$x[\bar{k}] = \frac{1}{N}\sum_{n=0}^{N-1} X[n] e^{\frac{2\pi i}{N} n(Np+k)} = \frac{1}{N}\sum_{n=0}^{N-1} X[n] e^{\frac{2\pi i}{N} nk} \underbrace{e^{\frac{2\pi i}{N} nNp}}_{=1} = x[k]. \tag{5}$$

Hence, writing the signal as in (4b) for arbitrary natural numbers k implicitly implies that it is periodic. Analogously, the DFT in (4a) is also periodic with period N.

Even though the implicit periodicity of the signal is indicated by the DFT definition alone, we still want to show the relationship between the DTFT and the DFT for periodic signals. As already pointed out, this does not allow for a rigorous mathematical treatment, since the DTFT does not exist. However, to circumvent this we shall make use of the standard convention that the DTFT of a periodic signal can be written in terms of Dirac δ-"functions". For an introduction to this concept, see e.g., [1] Appendix 2.A. More precisely, we shall make use of the DTFT of the constant function 1 expressed as (see [1], page 79),

$$\mathsf{DTFT}\{1\} = T\sum_{k=-\infty}^{\infty} e^{-i\omega kT} = 2\pi \sum_{n=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi n}{T}\right). \tag{6}$$

Now, consider the DTFT of the signal x, assumed to be periodic,

$$\begin{aligned} X_T(e^{i\omega T}) &= T\sum_{k=-\infty}^{\infty} x[k] e^{-i\omega kT} = T\sum_{l=-\infty}^{\infty}\sum_{k=lN}^{lN+N-1} x[k]e^{-i\omega kT} \\ &= T\sum_{l=-\infty}^{\infty}\sum_{j=0}^{N-1} x[lN+j]e^{-i\omega(lN+j)T} = T\sum_{j=0}^{N-1} x[j]e^{-i\omega jT}\sum_{l=-\infty}^{\infty} e^{-i\omega lNT}. \end{aligned} \tag{7}$$

For the second equality, we simply split the summation interval into blocks of length N. We then make a change of summation index ($j = k - lN$). Finally, we change the order of summation and make use of the assumed periodicity of x. Using (6) (with ω replaced by $N\omega$) we get

$$\begin{aligned} X_T(e^{i\omega T}) &= 2\pi \sum_{j=0}^{N-1} x[j]e^{-i\omega jT} \sum_{n=-\infty}^{\infty} \delta\left(N\omega - \frac{2\pi n}{T}\right) \\ &= \frac{2\pi}{N} \sum_{n=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi n}{NT}\right) \sum_{j=0}^{N-1} x[j]e^{-i\omega jT}. \end{aligned} \tag{8}$$

Here, we have made use of the scaling property $\delta(a\omega) = \delta(\omega)/|a|$. The presence of the δ-function turns the frequency continuous DTFT into a frequency discrete representation, by "cutting out" a set of discrete frequency values. Hence, we can replace ω by $2\pi n/NT$, yielding

$$\begin{aligned} X_T(e^{i\omega T}) &= \frac{2\pi}{N} \sum_{n=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi n}{NT}\right) \sum_{j=0}^{N-1} x[j]e^{-\frac{2\pi i}{NT} njT} \\ &= \frac{2\pi}{N} \sum_{n=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi n}{NT}\right) X[n], \end{aligned} \tag{9}$$

where we have made use of the DFT definition (4a) (periodically extended to all natural numbers n).

We can thus view the DTFT of a periodic signal as a frequency pulse train, with δ-functions spread on a discrete frequency grid $\omega = 2\pi n/NT$, $n \in \mathbb{N}$. Furthermore, each pulse is weighted with the DFT value at the corresponding frequency.

As a final remark of this section, there is an analogy to the relationship between the DTFT and the DFT for continuous-time signals. The DTFT, generally used for non-periodic signals, corresponds to the Fourier transform (FT) in continuous time. The DFT on the other hand, corresponds to the Fourier series (FS) in continuous time, and both apply to periodic signals. Just as in continuous time the periodicity of the signal will turn the frequency transform discrete, i.e., both the DFT and the FS are frequency discrete. Also, just as we have expressed the DTFT of a periodic signal as a Dirac sum weighted by the DFT, the FT of a continuous-time periodic signal is often expressed as a Dirac sum weighted by the FS coefficients.

4 Zero-padding - a time domain explanation

In the previous section we concluded that, when computing the DFT of a signal $x[n]$, $n = 0, \ldots, N-1$, we implicitly assume that the signal is periodic with period N. This insight can help us to understand how we can increase the frequency resolution by using zero-padding. Assume that we create a new signal by zero-padding x according to,

$$y[n] = \begin{cases} x[n], & n = 0, \ldots, N-1, \\ 0, & n = N, \ldots, M-1. \end{cases} \tag{10}$$

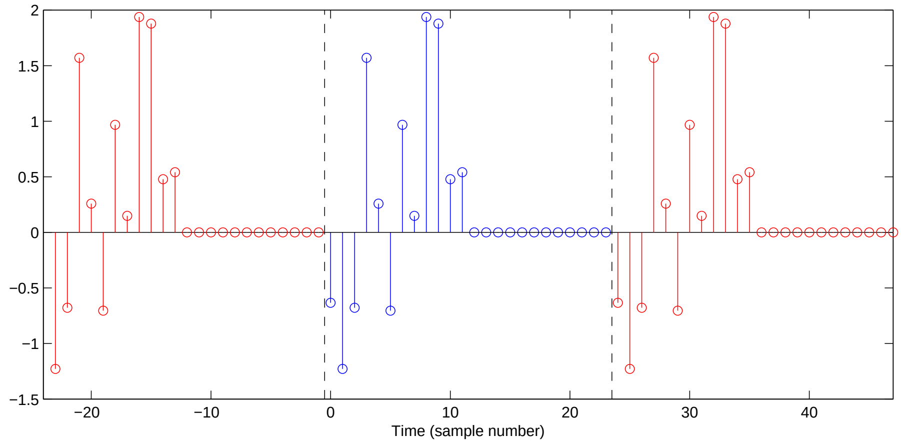

Figure 3: Zero-padded signal y, with the implicit assumption that it is periodic. Here, $M = 2N = 24$.

Hence, we simply augment x with $M - N$ zeros (of course $M > N$). From the definition of the truncated DTFT (3), it is clear that

$$Y_T^{(M)}(e^{i\omega T}) \equiv X_T^{(N)}(e^{i\omega T}). \tag{11}$$

However, it is not true that x and y have the same DTFT (only their truncated DTFTs coincide).

Now, if we compute the DFT of y, we make the implicit assumption that it is periodic with period M, as illustrated in Figure 3. As more and more zeros are appended to the signal, the periodicity assumption will more and more resemble the assumption that the signal is zero outside the sampled range. Hence, by padding the signal with zeros we "move" from the DFT assumption (periodicity) to the truncated DTFT assumption (that the signal is zero outside the known range). Consequently, the DFT of the signal will "move" toward the truncated DTFT, as illustrated in Figure 4.

Hence, zero-padding will indeed increase the frequency resolution. However, we do not gain any more information, we simply move from one assumption to another. If the (non-truncated) DTFT of x is thought of as the truth, i.e., what we really seek, then zero-padding will not necessarily be of any help. It all depends on which one of the two assumptions (periodicity or truncation) that most closely resembles the "true" signal.

With that being said, we should not understate the use of zero-padding, as it quite often can *reveal* (but not create) useful information about a signal. However, as a user it is important to understand what is implicitly done when we apply the tools of our toolbox. If we have this understanding, we are able to apply the tools only when it is appropriate to do so, and also to draw sound conclusions about the results that we obtain.

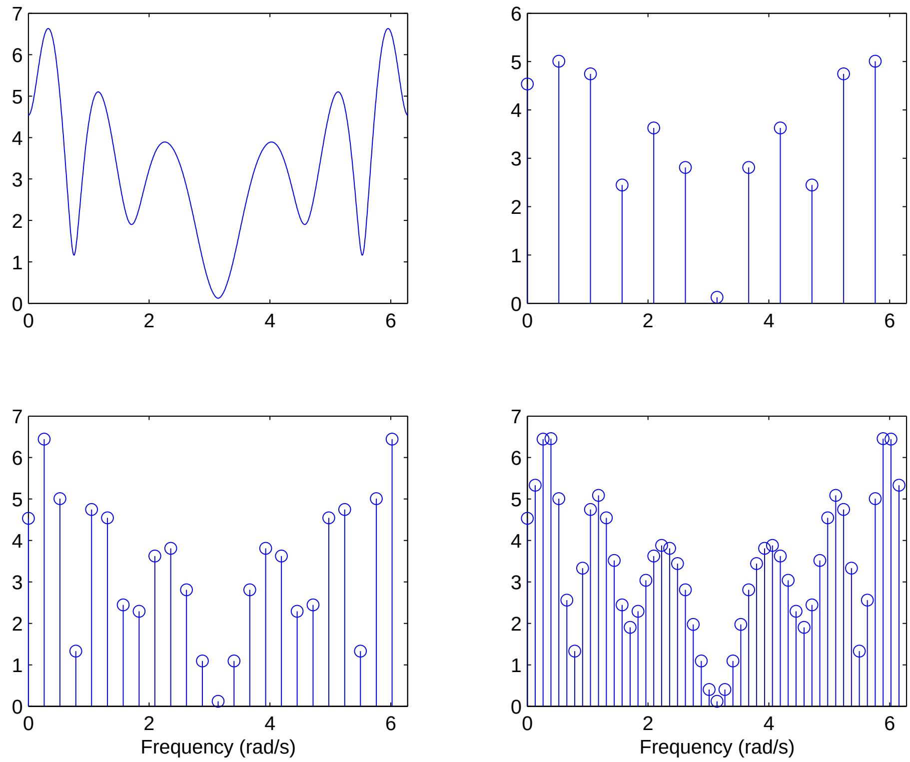

Figure 4: Zero-padding “moves” the DFT toward the truncated DTFT (the absolute values of the transforms are displayed in all plots). (Top left) Truncated DTFT of signal x. (Top right) DFT of signal x consisting of $N = 12$ samples. (Bottom left) DFT of the same signal, zero-padded to double length. (Bottom right) DFT of the same signal, zero-padded to three times the length.

References

[1] Fredrik Gustafsson, Lennart Ljung, and Mille Millnert. *Signal Processing*. Studentlitteratur, Lund, Sweden, 2010.